STATISTICS

STATISTICS

by

A. R. ILERSIC
M.Sc.(Econ)., B.Com., F.I.S.
Lecturer in Social Statistics at Bedford College (University of London),
Fellow of the Royal Statistical Society.

H.F.L. (PUBLISHERS) LTD
LONDON

Printed in Great Britain by
THE STELLAR PRESS LTD
UNION STREET BARNET HERTS

CHAPTER I

THE NATURE AND PURPOSE OF STATISTICS

According to a memorandum prepared by a committee of the Royal Statistical Society,[1] 'the science and methods of *statistics* in the modern sense range from the mere recording and tabulation of numerical data to subtle processes of inductive reasoning based on the mathematical theory of probability'. To allay any fears that may be aroused in the reader's mind, it can be stated that this study stops well before the latter stage. Statistics as defined in the latter part of the above quotation is a relatively recent development. Most of the significant advances in statistical techniques are little more than 50 years old, and recent years have witnessed a substantial expansion in the employment of statistical techniques in industry. The term statistician has until recently been very widely defined, the more so since there existed no recognised professional statistical qualification. In lieu thereof, graduates with degrees in mathematics or economics with some training in statistics formed the main source of statistical workers. Since the war, however, a new body, the Association of Incorporated Statisticians, has established such a qualification while the universities are extending the scope and scale of their teaching and granting degrees in which Statistics is the dominant, *i.e.* honours, subject. The undoubted importance of statistics is reflected in the large number of professional bodies which set a paper in statistics at some stage of their qualifying examinations. For this purpose, the student is required to learn the basic ideas and principles of the subject and it is, in the main, with such fundamentals that this book is concerned.

Definitions of 'Statistics'

The term 'statistics' is somewhat loosely employed to cover two separate concepts: (1) descriptive statistics; (2) statistical methods. 'Descriptive statistics' covers the collection and summaries of numerical data, popularly known as the 'facts'.

[1] Memorandum on Official Statistics. By a committee of the Council of the Royal Statistical Society, 1943.

Examples of this type of statistic are encountered every day in the national Press; one reads, for example, that the 'total working population of Great Britain at end February 1958 was some 24·1 millions of whom 16·2 millions were male and 7·9 millions female'. Figures relating to United Kingdom overseas trade and balance of payments appear regularly in the press. For example: 'total imports in March 1958 amounted to £320,959,138 compared with exports during that month of £294,210,055, so that for March there was an excess of total imports over total exports of £26,749,083'.

In the illustrations given above, the number of workers in Great Britain was given to the nearest hundred thousand, whereas the trade figures were stated to the nearest £. Actually, it would have been possible to have given the exact number of male and female workers in Great Britain as counted by the Ministry of Labour, just as it would have been quite permissible to give the trade figures in round millions. Two considerations determine the form in which such data are presented. First, and less important, is the purpose to which the figures are to be put; for simple comparisons it is often sufficient to state, *e.g.* 'in 1957-58 the United Kingdom government collected over £5,340 million of revenue as against £5,160 million in 1956-7'. The significant fact here is that tax revenues have risen substantially, not their precise yields to the nearest pound.

The second consideration is more important. How reliable are the figures themselves? Descriptive statistics are very often not exact or accurate in the arithmetic or accounting sense. Some data may be; for instance in Great Britain the number of Civil Servants and their salaries can be given precisely, but on the other hand no one can state *exactly* how many people there are in Great Britain at any date. This is true even for the census day itself, because errors due to omissions and miscountings are inevitable when enumerating some 50 million people living in nearly 14 million households. An exact figure is certainly not possible several years after that date, although the registration system, which should record every birth and death between the censuses, is very efficient. In any case, the figures for emigrants from and immigrants into the United Kingdom, are only estimates. However, since statistics is concerned mainly with large aggregates, exactness to the nearest ton, person, penny or any

other unit of measurement is, in most instances, quite unnecessary and, it must frankly be admitted, usually quite unattainable. The unemployment figures, for example, are given to the last unit, *i.e.* 472,618 persons were unemployed in February 1958. This is unquestionably an underestimate since a number of individuals temporarily unemployed would not yet have registered with the local labour exchange and would be excluded from this total. Consequently, the above statement normally reads '472,618 persons were *registered* as unemployed . . .', in other words, the figure is dependent on the fact of registration. Thus, while it is perfectly all right to give an exact figure in such cases, it should not be forgotten that the appearance of accuracy can be misleading, especially if for some reason the registration system is at all unreliable or temporarily affected by extraneous factors. It is far more important, when handling and studying statistical data, to ensure that the source of the original data is reliable and that the data collected are relevant to the question to which an answer is sought. Whenever any statistics are consulted, the first question is not, 'To what unit or place of decimals are these figures reliable?' but 'Where were the figures obtained, by whom, and why?' If satisfactory answers to these questions can be given, few qualms as to the reliability of that data need be entertained.

Statistical data in their raw state are by themselves of little value. It may be interesting to the managing director of a company to learn that the sales of a particular product in one country total £500,000 last calendar month, but the information only becomes statistically significant when related, for instance, to the fact that a turnover of only £300,000 was achieved in the corresponding month of the preceding year. The essence of statistics is not mere counting, but *comparison* – and provided this fact is always borne in mind, careless and consequently useless compilation of data may be avoided. Comparisons are valid only when they are between quantities expressed in identical terms, *e.g.* a direct comparison of weekly earnings as an indication of living standards between two periods is invalidated if in the first the working week was 45 hours and in the second only 40 hours.

A good example of careless and consequently irrelevant comparison was provided by a Board of Trade Press notice concerning the tobacco shortage in 1949. After pointing out that

more tobacco was reaching the shops, it continued: 'Against this is the fact that more persons are smoking than before the war, the United Kingdom population having risen from 47,700,000 in 1939 to 50,000,000 at the end of 1949'. Since the population can only expand by immigration and by new births, and the former was relatively insignificant in this period, it might be assumed that the infants were tobacco addicts at an early age! The relevant figures were, of course, those of the *adult* population which, due to the changing age structure of the whole population, has increased.[1]

On the other hand, too much should not be read into a comparison of such data. The results should be scrutinised because there is a difference. The question to be asked is then not so much 'how big is the difference', but 'why is there a difference'. The following illustrations may bring home the point. An examination of the marriage rates in the United Kingdom show occasional but sharp fluctuations between the relative numbers of marriages registered in the first and second quarters of the calendar year. Such fluctuations are largely attributable to the fact that the date of Easter varies and occasionally falls within the first quarter, although a new influence in recent years has been the attraction of an income tax refund if the marriage takes place before 6th April. Comparisons of quarterly births in the United Kingdom have not always been straightforward. After 1941 the number of births registered in the final and first quarters of successive years had to be adjusted to provide a true comparison with previous years. Previously, births were often registered up to six weeks after the event, as allowed by law. In consequence December births were often registered in January, thereby inflating the following quarter's total. But, with the introduction of ration books, the arrival of a new baby – subject to registration – meant a new and extra ration book. The elimination of the delay between the two events markedly affected the official statistics, yet at the time there was a very simple explanation of the sudden change in the birth pattern! Similarly, the sharp increase in the number of divorce petitions in the immediate post-war years gave rise to much public comment and suggestions that family life was disintegrating. Further reflections indicated that although the number of petitions had greatly

[1] Even this needs qualification, since it assumes that the smoking habits of the average adults are unchanged, or at least that the *per capita* consumption has not decreased.

increased, the increase could be explained largely by the fact that

1. The war had led to an accumulation of petitions which would normally have been spread over the war-years thus avoiding the marked post-war accumulation;
2. That the enforced separation of husbands and wives during the war inevitably led to what now seems to be a once-and-for-all increase and the cumulative effect of wartime influences was only apparent with the return of peace.

Thus the collection of statistical data and their verification regarding source and definitions used, forms only the first part of the statistician's work. The statistics themselves prove nothing; nor are they at any time a substitute for logical thinking. There are, as pointed out above, many simple but not always obvious snags in the data to contend with. Variations in even the simplest of figures may conceal a compound of influences which have to be taken into account before any conclusions are drawn from the data.

Statistical Methods

As indicated in the quotation in the opening paragraph, statistical methods range from simple numerical processes to 'subtle processes of inductive reasoning based on the mathematical theory of probability'. These processes are also covered by the term *statistics*. but to distinguish this meaning of the term from our first definition – descriptive statistics – they are usually referred to as *statistical method.* The simpler methods which involve little more than elementary arithmetic are discussed in the chapters which follow. The really fascinating developments in technique, which have made statistical science one of the most powerful tools in the hands of the modern research worker, are of relatively recent origin, mainly in the last thirty years. In industry, particularly since the war, statistical methods have been extensively employed to control the quality of the product and to ensure that faulty goods are not sent out. A brief introduction to this technique, known as *quality control*, is given later.[1] Another important development has been that dealing with the design of experiments and their subsequent analysis, which is of course determined by the actual design. Modern

[1] Chapter XX.

designs are much more efficient in that they enable much more information to be extracted from a given amount of data. These techniques, however, require a mathematical ability of no mean order and are only mentioned in this text to emphasise the potentialities of statistical method.

Perhaps the most useful and generally applied technique devised by the statistician is known as *sampling*. Most readers will have read the election forecasts of the public opinion polls; many will have heard about market research whereby manufacturers learn what the consumer thinks of his product by posing specially designed questions to a selected few people. Fortunately, although this branch of statistical method is complex in application, the lay reader can acquire an understanding of the principles underlying sampling techniques which will enable him to appreciate what the statistician terms the 'significance' of his results from an analysis of sample data.[1] Information and conclusions based upon sample data, although they are reported with all their limitations are later often misquoted and bandied about with a complete disregard for their limitations which the statistician has so carefully emphasised. It is hardly surprising that the statistician so often hears the dictum, 'lies, damned lies and statistics' quoted at him. Yet it is not the statistician who is at fault; it is the individual who quotes – often selectively to support his own tenuous arguments – the statistical results without any of the qualifications with which the statistician has surrounded his conclusions. Some of the problems of statistical inference, in other words deducing information from a limited amount of data, *i.e.* a sample, are discussed in non-technical terms later.[2]

Statistics in Business

Every business man today appreciates the value of accurate and regular financial statements in the conduct of his business. The widespread expansion in management accounting in recent years is attributable to the growing realisation on the part of the industrial community that without facts concerning output, costs, turnover, expenses, etc., it is impossible to conduct the affairs of a business so as to obtain maximum efficiency.

[1] Explained in Chapter X.

[2] Chapters IX and X.

Whatever the legal form of a business, be it public or private, a public corporation or a unit of nationalised industry, management today has become highly complex. The administration must make its decisions in the light of facts prepared by the executives rather than – as was the case in the smaller businesses of the past – from intimate personal knowledge of the concern. Although no volume of statistics can replace the knowledge and experience of the executives, it supplements it with more precise facts than were hitherto available. The value of statistics in the large concern is indicated by the comments of Lord Heyworth, Chairman of Unilever, in his Presidential Address to the members of the Royal Statistical Society[1]: 'We also have an interest in statistics of a general nature. We operate in most of the countries of the world. We frequently have to decide whether it would be, say, a better proposition to erect a new factory in Malaya or extend our ice-cream business in England. To come to a satisfactory decision we have to know all about not only the specific conditions of the trade in which we are proposing to engage in Malaya and England, but also the whole general health of their economies . . . So for such purposes a whole range of outside statistics may be relevant, from the amount of debt in the countryside to the number of votes obtained by communists in municipal elections.'

It is much easier and probably more important for the larger undertaking to assemble and prepare statistical data relevant to their problems. This does not mean, however, that the smaller firm can derive no benefit from the simpler statistical techniques outlined in later pages. The mere assembly of data relating to the financial and production activities of the firm in tabular form and simple charts can often bring to light facts and trends which were hitherto not fully appreciated or not apparent from a scrutiny of the bare figures themselves. Given some knowledge of even the simpler methods, errors in the interpretation of such statistical data can often be avoided.

Finally, in government publications there is a wealth of information relating to virtually every industry as well as to the economy at large, accessible to all. One government committee found that in contrast with American business, British industrialists and businessmen made little use of such data as are

[1] 'The Use of Statistics in Business' – Lord Heyworth: Presidential Address to Fellows of the Royal Statistical Society, read 31st October, 1949.

compiled from the periodic censuses of production and distribution. Yet it is from such information that the business man can make a better assessment of the economic situation when forecasting the prospects of his business and provide him with information relating to other problems directly concerning him; for example, prospective supplies of labour, raw materials, machinery, building, etc. To ignore such information is at best unwise and at worst may be likened to a ship's captain who ignores his barometer in assessing the weather prospects.

Government Statistics

Most of these data are the by-product of government activity in the social and economic spheres which has brought with it the need for positive and constructive legislation in economic affairs. The acceptance by all parties of responsibility for the maintenance of economic stability, which was acknowledged in the White Paper on Full Employment[1], necessitates the collection and preparation by the responsible Government departments of statistics covering the whole vast complex of the nation's economy. Many of these data can only be obtained from the business and industrial community, and this involves the generally disliked task of form filling. There is scope for much education on both sides in this matter. The Government statisticians responsible for the preparation of the forms and the presentation of the results will perform their tasks more efficiently given reasonable co-operation from their informants.

Statistical techniques enable the Government Actuary to estimate the potential demand for pensions, sickness benefit and unemployment allowances. The B.B.C. has, since October 1936, maintained a statistical department for audience research in order that programmes should as nearly as possible meet the demands of the public.[2] The Budget, which influences every side of the economic life of the community, is itself based on estimates of the probable consumption of taxed commodities, prospective income levels and mortality rates (for estate duty receipts), all of which owe their accuracy to statistical techniques. There is little doubt as to the importance attached by the government to reliable and up-to-date statistical data on the economic scene.

[1] Employment Policy. Cmd. 6527.

[2] *Methods of Listener Research employed by the B.B.C.* – R. J. E. Silvey, J.R.S.S., Vol. CVII pts. III/IV, 1944, p. 190. A brief description of this work is given in Chapter XIII.

The Central Statistical Office is expanding continuously its output of economically significant series such as hire-purchase statistics and estimates of investment outlays by large firms in the United Kingdom. Much is being done to improve the economic statistics needed for budgetary policy. Mr Macmillan, when Chancellor, likened the then available data to 'looking up the trains in last year's Bradshaw'.[1] Economic planning is impossible without accurate statistics published with the minimum of delay.

The contents of the following chapters may at times appear to be far removed from these important and fascinating problems; but just as a mathematician has first to learn the multiplication tables, so the potential statistician must first acquaint himself with the elementary principles of statistical analysis. If after finishing this book, the reader is still interested in the subject, then – his mathematics permitting – a brief bibliography offers him additional material for study.

[1] Budget address 1957. Many of the more important economic statistics are discussed in Chapters XIV and XVII.

CHAPTER II

STATISTICAL DATA: DEFINITIONS AND SOURCES

Statistical method consists of two main operations; counting and analysis. The analysis may entail no more than a simple comparison, *e.g.* at last Saturday's football matches there were 60,000 spectators at Chelsea and 40,000 at Tottenham. From these data at least one conclusion can be drawn, *i.e.* there were 50 per cent more spectators at one match than at the other. The ensuing question, 'why this was so?' requires more information than is given here. At more advanced levels statistical analysis is very much more complicated and is based on specially designed mathematical techniques. But to whatever type of analytical technique the data are to be subjected, the first stage of any statistical enquiry is the collection of the facts and this means counting. The statistician has no use for information that cannot be expressed numerically, nor generally speaking, is he interested in isolated events or examples. The term 'data' is itself plural and the statistician is concerned with the analysis of aggregates.

The data themselves may be of any kind as long as they can be counted. They must, however, be susceptible to classification as well as to counting. In the statement that there are 50 million people in Great Britain, the unit of counting is a person, just as in the figure of unemployed quoted earlier, the unit is a person registered at the local Employment Exchange as unemployed. The actual process of counting, whether it be the census enumeration of heads of households and the persons in them, or the number of insurance cards lodged at the Employment Exchange on a given day each month, is easy enough, although with very large aggregates mistakes will occur. But the latter are, generally speaking, insignificant in relation to the absolute figures in the totals. A much more serious consideration for the statistician is to be certain that not merely have all the units been counted, but that only the units relevant to his enquiry are

included. For example, the characteristic of the unit in a count of the unemployed is the fact of being both unemployed and registered as such. The Ministry of Labour publishes, among other analyses, totals sub-divided by sex and age (actually as between men and boys, women and girls) by the duration of unemployment as well as the location. All these analyses depend on classifications which in turn depend on the definitions employed. Clearly, some definitions are easier than others, *e.g.* male and female, men and youths – the latter being defined as under 18 years of age.

Some classifications, however, are more arbitrary since the definition of the unit is itself arbitrary. For example, suppose we sought to classify a group of women by the colour of their hair or eyes. Inevitably there will be some colours which will be difficult to determine although if the same women were classified by height, the classification would be simple. The basis of classification depends on the nature of the characteristic by which units are being identified and counted, *e.g.* women over 16 years of age in the United Kingdom are either married, single, widowed, separated or divorced. The characteristic in this case is described as an *attribute*. The same women can be classified by age; in this case the classification is based on a characteristic which is known as a *variable*. The distinction between 'attribute' and 'variable' is important but quite simple. The former describes a characteristic which is not capable of numerical definition, *e.g.* colour of eyes, houses condemned as 'unfit for habitation', recruits classified as 'grade one'. A variable is a characteristic which can be expressed in quantitative terms, *e.g.* height in inches, salary in pounds, or marks in an examination.

The point of the foregoing paragraph is to emphasise the fact that the statistician is usually concerned with groups and 'populations' consisting of units possessing a common characteristic, although he may compare such groups which are dissimilar in one particular respect. For example, in the pre-election polls it is usual to classify the respondents (*i.e.* those who have been interviewed) by their political party to ascertain how far allegiance to a party affects their attitudes to particular problems of the day. But, it should be noted, all the units in the various groups share a common characteristic, *i.e.* the right to vote, and it is for this reason that they have been interviewed.

The significance of precise definition becomes apparent when information is being collected. This is especially true of what are termed 'secondary' statistics, *i.e.* statistical data available in published form such as the Annual Abstract of Statistics.[1] Especial care must be taken in using published data to ensure that over the relevant period of time the definitions or coverage of a series of data have not been altered. For example, it is impossible to compare pre- and post-war living costs by simply comparing the pre-1947 Cost of Living Index with the current Index of Retail Prices. Similarly, the published official estimates of the 'working population' at the present time are compiled in a different way from the figures published at the end of the war so that comparisons of 'total working populations' in 1947 and 1957 can only be made after adjusting certain figures to a common basis.

Such considerations are especially relevant when it is learnt that by far the most important and prolific source of published statistical data is the government. Each department produces a great deal of statistical data, primarily as a by-product of its administrative functions. For example, the Home Office is responsible for the preparation of the Annual Report on 'Criminal Statistics' which gives information on the extent and type of crimes committed, the activities of the police in clearing up such crimes and the punishment meted out to the convicted offenders by the Courts. The published statistics on crime are especially interesting in view of their chequered history.[2] For example, the only crimes which officially exist are those known to the police. Thus, it is probable that the extent of blackmail and sexual assault is understated in the annual returns, since the victims are often unwilling to run the risk of publicity if they were to charge the offender. Fluctuations from year to year in the published figures may reflect not only changes in the incidence of a particular crime, but mainly the fact that the police have at intervals undertaken a major drive against it. This, for example, probably explains the apparent increase in homosexual offences of recent years. It is unlikely that the extent of such practices among the male population has changed as much as recent publicity given to court cases would suggest. The increased number of convictions merely indicates that the police

[1] A compendium of official statistics compiled by the Central Statistical Office and published annually. The Abstract for 1958 is the 95th in the series.

[2] See Chapter XIV for references.

have become more active in trying to suppress these practices. Contrasting rates of crime, *i.e.* of certain types of offence in different areas of the country, may merely reflect local police directives or the prejudices of the local bench of magistrates. For example, where the bench is reputed to be lenient towards certain types of offence, it is probable that the local police will be more reluctant to charge the offender than in another area where the bench is not so inclined.

In brief, the greatest care must be exercised in using any statistical data, especially when it has been collected by another agency. At all times, the statistician who uses published data must ask himself, how were the data collected, by whom and for what purpose? Many official statistics, *i.e.* government produced, are reliable and complete. But this is not true of all published data, least of all that covering a long period of time. While there has been a great improvement in the quality of official statistics since the last war, it is not so many years since the 'working population' was differently classified by the Board of Trade, the Registrar-General and the Ministry of Labour, thereby making comparisons of similar data collected by these three government departments virtually impossible.

With these provisos and warnings regarding the use of published data, some indication of the scope and extent of published statistics can be given. Published sources can be divided into two categories. The *first* and most important are the official statistics prepared by the government. These are of two kinds: those which are the administrative by-product of a department's daily work and those which are collected at intervals for specific purposes. An example of the first type are the figures prepared by the Customs and Excise department. The bulk of these appear in the annual report of the department. Some of these annual reports, *i.e.* Blue Books, contain a wide range of statistical information. For example, apart from such routine statistics as the amount of tax collected and number of assessments made, the annual report of H.M. Commissioners of Inland Revenue has been considerably expanded in recent years and is now a mine of information regarding the size and distribution of incomes and direct taxation. Examples of the second type of official statistics are the various censuses, *e.g.* the decennial population census and the census of production. Apart from information of a

statistical character contained in Blue Books and special census reports, there are the periodic White Papers such as those on the Balance of Payments and the Economic Survey. These papers are usually prepared by the appropriate department, but since the war there has been created a Central Statistical Office which has two main functions. First, it is responsible for co-ordinating the statistical information produced by the various departments so that, for example, with a Standard Industrial Classification introduced in 1948, all departments now prepare data classified (where this is appropriate) under the same industrial headings and not as before the war, as was pointed out above, on their own classification.[1] Second, the C.S.O. as it is usually termed, is responsible for bringing together the economic statistics needed for the formulation of policies designed to maintain economic stability. Many of them are published in the *Monthly Digest of Statistics* and *Economic Trends*. The more important of the United Kingdom economic statistics are discussed more fully in Chapter XVII.

The *second* category of published statistics is not so large, nor is it so well known. Local authorities publish annually a wide range of statistical information relating to their financial and social activities, as do the various nationalised industries, *e.g.* the National Coal Board. Some of the City institutions produce statistical studies of matters which are of especial interest to them. For example, the Midland Bank has long published a series of statistics showing the amount of capital raised in the capital market by governments and companies, while Lloyds Bank publishes an index of the supply of money. In recent years the need for the independent collection of economic statistics has been greatly reduced by the government's recognition of its responsibilities in this field. There remains, however, one branch of statistical work of the utmost importance for the private investigator, either individual or, more usually, corporate. This is the carrying out of surveys among special groups of the population to obtain information which is not otherwise available. Thus the Department of Applied Economics in Cambridge has conducted local surveys into family budgets and the Oxford Institute of Statistics was responsible among other surveys for a pioneer enquiry into personal savings in the United Kingdom.

[1] The S.I.C. has been amended slightly in 1958.

Another private organisation conducted a national enquiry into the conditions of life among the 'over-seventies', while a professional body annually carries out sample enquiries into the reading habits of the community. Although initiated and developed by non-government bodies, the technique of sample enquiries has been adopted by the government since the war. It has its own survey organisation known as the Social Survey and large scale sample surveys have formed the basis of the Retail Price Index which serves as a cost of living index in the United Kingdom. Using similar methods, there is a continuous survey into the food consumption habits of the community, the findings of which are published in the annual reports of the National Food Survey.

If it is proposed to use statistical data which have been obtained from a survey carried out by some other party, then it is imperative that the report of that survey be read, special attention being paid both to the sample design, *i.e.* the informants and whether they were representative of the larger population, and to the questionnaire. Any survey report which lays claim to serious attention will give the reader such information within its covers.

Similar considerations arise when the statistician is forced to collect his own data direct from a group of respondents. Questions must be carefully phrased so that they are unambiguous and mean the same to all respondents. It will often be necessary to define certain terms. For example, 'household income' covers all wages and salaries before tax, as well as pensions and tax-free spare-time earnings. Individuals in the Household Expenditure Survey of 1953-4 often gave their net incomes after deduction of tax and national insurance contribution, as well as omitting any cash receipts from spare time activities.

The importance of paying attention to the precise definition of statistical units and sources of data collected by others becomes evident when it is recalled that the conclusions derived from the enquiry can only be as good as the original data upon which they are based. You cannot make, runs the old adage, a silk purse out of a sow's ear. It is equally true that an ill-classified collection of inaccurately defined data from misguided informants will not provide reliable data for a useful statistical study. It is for such reasons as these that the Central Statistical Office has prepared a

small booklet of terms and definitions of the units and data published in the *Monthly Digest of Statistics.*[1] The student would be well advised to look through this, not for the purpose of learning any of the definitions, but first so that he becomes aware of the care which must be devoted to this apparently elementary branch of statistics, and second, to ensure that when he abstracts figures from the *Monthly Digest* he will do it properly. He will also understand why it has been so necessary to devote an entire chapter to what appear to be commonsense propositions! If he is a reader of crime-fiction, he may recall the comment of so many detectives, 'in the long run, it's the apparently small and unimportant things that count'. This is just as true of statistical source material.

[1] This Supplement on Notes and definitions is published annually in January by H.M. Stationery Office.

CHAPTER III

COLLECTING THE DATA

Defining the Enquiry

As was indicated in the preceding chapter, there is no lack of published statistics on a wide range of subjects. The usual difficulty, except in the case of the better known sources, is to locate them. It is good practice for any statistician to assume that whatever information he himself may be seeking, someone else has had the same idea before him and has probably done some work in the field. Even if what has been done does not prove very helpful, it is a poor statistician who cannot profit to some extent by another field worker's or researcher's work, or even his mistakes. In practice, however, the difficulties of actually collecting data (whether from published sources or by means of an enquiry) are often not as great as the difficulties which arise at the outset in defining the scope of the enquiry. There is first the problem of deciding exactly what information the statistician wants; then the best way of getting it. For example, the managing director of a large works has decided that there is 'too much absenteeism' among the labour force and you are instructed to look into the matter. Where does one start?

First one has to find out how many workers have lost time during a given period. But what period? In winter absences due to legitimate sickness are likely to be high; in summer the extent of absenteeism may be lower than usual because with a holiday in the offing the worker will not want to lose money. Assuming a period for study has been selected, how long should it be? A month, three months or longer? Having decided these points, how does one find out from the worker why he lost time? If a large number were away on the day when nearby football teams were meeting in a local derby, the absenteeism may easily be explained, but it is unlikely that when asked the worker will give the real reason for his absence. Are we then to assume that unless the absence can be justified by the production of a medical certificate, it is 'absenteeism'. What about the family man who

loses a morning's work to look after the family when the wife is sick? A good excuse, perhaps, but one which is difficult to refute. The reader may like to ponder on this particular illustration to see if he can devise a solution!

When the enquiry is on a broader basis than a purely localised survey, the problem arises as to the best method of contacting those whom one hopes will give the information needed. How can this best be done? To visit and interrogate them all personally might be desirable, but quite apart from expense it may be impossible on account of time. If one writes to them and sends a list of questions, how does one avoid the considerable risk that the form and letter will go into the waste-paper basket unanswered? Not least, as the reader will doubtless appreciate from personal experience, there are many questions he would like to ask at times, but the chances of getting either a polite or truthful answer are slight. If they cannot be asked directly, is there some other way of getting an indication of the facts? For example, an enquiry into health could include the question 'have you during the past month visited your doctor?' The object of the question is to learn whether the informant has been unwell during that period, but he may have felt unwell and still not taken medical advice. Consequently further questions are needed to build up the full story. The danger of this approach is that one finally finishes up with a schedule of questions which will take at least an hour to ask, so that informants will then refuse to be interviewed! However, sufficient has been said to indicate briefly some of the statistician's problems in the survey field. The techniques of surveys and interviews are discussed in detail later.[1]

Collection of Data

There are three basic methods of collecting the information needed:

1. The investigator may interview personally everyone who is in a position to supply the information he requires. Such a procedure will be possible in very few cases indeed, since most statistical enquiries cover a wider field than any single investigator could possibly examine personally within any reasonable time. An interesting example of such an enquiry is

[1] See Chapter XIII.

that conducted by Professor Zweig who personally interviewed 400 people.[1]

2. The task of interviewing informants may be delegated to selected agents who will be provided with a standardised questionnaire and explicit instructions as to the mode of its completion and the information to be elicited. The main problems in this case are the selection of suitable agents and the cost involved. It is not merely sufficient that they should be given routine instructions; they should be fully conversant with the purpose of the enquiry, since inefficient interviewing will seriously affect the value of the results obtained. Against the disadvantage of high cost and the difficulties of obtaining suitable agents must be set the very considerable advantage that the information received will probably be highly reliable. Such interviewers are nowadays employed by all research organisations such as the Social Survey and the market research offices.
3. The last method is by questionnaire addressed to individual informants. This method, at one time extensively employed, possesses the apparent advantage that a very large field of enquiry may be covered at relatively low cost, and the larger the coverage the less significant will be occasional errors in the filling up of individual forms. This method of collecting information is not very satisfactory due to the low proportion of returns. The government uses it for various census enquiries, *e.g.*, population, election, production, but in these cases the return of the form is compulsory. If it is voluntary, only those individuals, generally speaking, who are particularly interested in the subject matter of the enquiry will trouble to return the questionnaire. Then the investigator may merely have a collection of biased data.[2]

Drafting the Questionnaire

It is idle to pretend that mistakes are not made in completing a schedule of questions. More usually due to carelessness, they may sometimes be deliberate, *e.g.*, the over-statement by a manufacturer of the proportion of total output allocated to

[1] *Labour, Life and Poverty*, F. Zweig, Gollancz, 1948.

[2] See Chapter XIII for a discussion of postal enquiries.

export as against the home market in a Ministry of Labour return, or they may be accidental, made in all good faith, but due to a misunderstanding of a particular question. In any enquiry the compilation of the questionnaire is difficult, and it is often the joint product of several specialists in survey work. It has been justly observed that a survey is no better than its questionnaire. This is especially true where the results of the enquiry depend on the goodwill and voluntary co-operation of the prospective informants.

If a questionnaire sent by post is unduly long, of too prying a nature, or the questions are too complex, it is quite likely to be consigned to the wastepaper basket. Similar weaknesses in a schedule used by an interviewer will lead to a high refusal rate. It may be that some of the copies returned will be inaccurately completed owing to a badly-phrased question. Such information is then valueless for further statistical analysis, since it is not possible to tell whether the informant was answering the question as set, or putting an altogether different interpretation on it. In a public opinion survey the question 'Do you believe in God?' was put to the respondent; from a statistical point of view the only justifiable conclusion that could be drawn was that *X* per cent. of those questioned *said* they did. What they meant, or what they understood by the question, only the individual informants will ever know. A paramount consideration, when preparing the questionnaire, is the type of individual to whom it is addressed. The population census, which is compulsory, goes to all heads of households in the country, and since the mental ability of some fourteen million informants will vary considerably, the questions must be worded so that the least intelligent individual cannot fail to grasp their meaning. On the other hand, a questionnaire sent to business executives may contain questions of a technical nature. Nevertheless, clarity and simplicity in their construction are still essential. Whenever the postal method is used, the questionnaire should be accompanied by a short letter explaining not only the purpose of the enquiry, but also the advantages which will accrue to the informant from the final results. There is often no finer goad to co-operation than self-interest!

Since the value of the results obtained from the enquiry depends largely on the adequacy of the questionnaire, the following

points should be borne continuously in mind during its preparation, even in cases where it is to be completed by an official agent or the investigator himself:

1. Few people enjoy form-filling or answering questions. Keep it as short as possible.
2. Complicated and long-winded questions sometimes irritate and often confuse the respondent and result in careless replies. Make the individual questions short and simple.
3. Answers such as 'Probably', 'Fairly good', 'Average', mean nothing to a statistician, since they signify different degrees to different individuals. Ensure that all questions may be answered as far as possible by either 'Yes' or 'No', or by a name or figure. Some readers may, however, have encountered the type of questionnaire circulated by the Audience Research section of the B.B.C., in which the listener is asked to indicate whether he or she considered a particular programme excellent, good, fair, or poor. This survey is concerned with estimating the number of listeners who 'thought' the programme was good, *i.e.*, the listener's *subjective* assessment of the programme. If 90 per cent. of the listening public answered 'excellent', then as far as the B.B.C. producer is concerned, the programme was outstandingly successful. There is no attempt here to arrive at an impartial, *objective* assessment of the quality of the programme, since it is clearly impossible to do so. In short, the questions and their answers depend on the purpose to be served by the enquiry.
4. The questions should follow a logical sequence, so that a natural and spontaneous reply to each is induced. Thus it is clearly politic to enquire whether a woman informant is married before asking her how many children she has! Similarly, if the enquirer wishes to know the amount spent on children's clothing, a comment in passing on the tendency of children to 'kick their shoes to pieces' may evoke for the trained interviewer a more adequate and reliable response than a direct question.
5. Few people willingly provide intimate facts about themselves. Many resent such questions, which should be avoided as far as possible. In some cases where private information of this nature is needed, the method of personal enquiry may be most likely to yield results. Such questions should, however,

be kept to the end of the interview, when the informant may feel more at ease with the interviewer.[1]

6. When public opinion on a particular issue is being assessed, it is important to ascertain from the respondent whether he has any knowledge of the subject, before asking his opinion on it! Opinion polls, as compared with factual enquiries, give rise to a host of very complicated problems in questionnaire design, some of which are discussed in Chapter XIII.

It need hardly be added that once the questionnaire has been drafted, it should first be tested on a small number of individuals, to assess the probable reaction of the wider public to be covered later. This policy was pursued by the Board of Trade in the 1948 Pilot Census of Distribution. It was proposed to carry out a census of shops and service establishments in 1951, but the Department undertook initially a sample survey of a few selected areas and trades representing the whole country. The lessons drawn from the many criticisms and suggestions made enabled improvements in the questionnaire to be introduced, thereby facilitating the task of the informant and ensuring more accurate and prompt replies.[2] The mere fact that the census is enforced legally does not help the statistician very much; he is much more anxious to obtain a fair sample of accurate replies than a mass of carelessly compiled and often inaccurate information.

Technique of Sampling

As already stated above, it is frequently physically impossible to obtain a really wide coverage of information from all those individuals who might come within the scope of the enquiry. This is especially true where, owing to the nature of the enquiry or survey, it has been decided to use agents interviewing their subjects personally. Examples of such surveys are the 1953/4 Ministry of Labour enquiry into some 12,000 'working-class' household budgets, the Ministry of Labour enquiry into the attitudes of retired workers to retirement, the many surveys carried out by the Social Survey on behalf of various Government Departments, and finally the well-publicised political polls

[1] The extent to which people will supply information even of the most intimate nature, is emphasised by the Enquiry into Fertility. See Royal Commission on Population's 'Family Limitation,' R.C.'s Papers, Volume I. Admittedly in this case the information was collected by doctors and midwives. Most surveys include questions relating to the family circumstances and respondents are usually prepared to answer them.

[2] In view of the widespread protests and even hostility encountered with the Pilot survey, it is clear that as a result of the lessons learnt, a great deal of expense was avoided.

which are discussed regularly in the daily press. In these cases only a small part of the field may be covered; estimates of national public opinion are often based on a sample of less than 2,000 informants. Yet it is true to say that in most cases, provided that the survey has been conducted with due regard for statistical technique and principles, the results should differ only to an insignificant extent from those which would be derived from a complete census.[1] The validity of this contention has been proved in the past by careful checking, repeated sample enquiries, and in some cases by a full-scale enquiry where the field to be covered was not impossibly large.

This system of selection of a part of the whole, known in statistical method as 'sampling', may be explained by a simple analogy. The practice of sampling is frequently encountered in the world of commerce. Thus a small sample of tea or grain, of cloth, or of many other commodities, is frequently the only means whereby a prospective buyer can assess the quality of the bulk. The principle underlying the process is the assumption, generally borne out in practice, that the part is genuinely representative of the whole; thus the Public Analyst bases his report on the quality of a product on a few tested samples, and the assayer assesses the mineral content of an area on samples of ore selected at random. Innumerable experiments have revealed that, given two conditions, a selection from a large group of individual items will possess the characteristics of that group. The two conditions are: firstly, the selection of the sample must be random, *i.e.*, each individual item within the aggregate must have an equal chance with any other of being selected. Unless this fundamental requirement of any sampling technique is most carefully observed, the sample will probably be biased and as such virtually useless. Secondly, the samples selected must not be so small that all the possible varieties within the group cannot be duly represented. If a large sack of marbles contains ten colours, it is not sufficient merely to extract enough marbles to include all the ten colours, the number extracted must be adequate to permit the sample to contain proportionately as many reds compared to, say, blues, as there are in the whole sack.

It follows as a matter of course that the larger the sample the more closely the final results will approximate to those which

[1] A complete enumeration of the entire field is known as a 'census.' A limited enquiry based on part of the field is usually described as a 'sample survey.' These are discussed more fully in Chapter XIII.

would be derived from a census, although after a certain stage the additional precision in the sample result gained by increasing the sample may be small.[1] 'Precision' in this context means the likelihood that the sample result will not differ to any appreciable extent from the result to be expected if a complete enumeration had been carried out. The proportion that the sample bears to the whole field (usually known as the 'universe' or 'population') is unimportant in determining the degree of precision. It is the absolute size of the sample that is relevant.[2] In recent years statisticians have devoted considerable thought to the task of ensuring the representativeness of the sample and in determining the minimum sample size required to produce a reliable answer to their problems. Some of the main sampling techniques used in surveys are explained later in this book.[3]

Problems in Sampling

Where the data are not completely homogeneous, *i.e.*, comparable by virtue of a single common characteristic, the sample must be carefully selected. Thus, in a survey designed to ascertain the average expenditure on food of all types of individual households, there is a twofold problem. Firstly, a 'wealthy' family will spend more money and buy different commodities than a 'middle-class' family, while the latter will probably spend more than a 'working-class' family. Merely to aggregate all three types and extract an average outlay per family would be virtually meaningless, since the three groups are not comparable for this purpose. Thus separate samples from each of the three groups must be taken and compared independently. This problem was encountered in the Household Expenditure Enquiry 1953-4, the data from which were used to construct the Index of Retail Prices. By excluding the 'well-to-do' households and the very poor pensioner, the remaining households formed a homogeneous group in respect of their spending habits.[4] If they are to be combined to answer the different question of how much an average family spends, then the result for each social class must be included in due proportion as these classes exist in the

[1] See Chapter X.

[2] See Chapter X.

[3] See Chapter XI.

[4] This subject is discussed more fully in Chapter XIV

whole population. This process, known as *weighting* is discussed in Chapter VI.

Secondly, for example in a public opinion poll which is designed to reflect the views of the whole adult population, the individuals interviewed must not only represent each section of the community, *e.g.*, rich and poor, professional workers and wage earners, etc., but the numbers sampled from each group must bear the same relationship to each other as do the separate groups to the whole population. Non-observance of this simple rule resulted in the bad failure of an American journal's forecast of the 1936 presidential election result. In that survey too large a proportion of the wealthier and professional workers were included in the total sample. The more recent failure in the 1949 U.S. presidential election of the public opinion poll forecasts can perhaps be attributed in part to the fact that the proportions of the *voters* from each social group were different from the actual proportions of the *electorate* in each social group. Thus, if 60 per cent. of the electorate voted, 80 per cent. of the 'working-class' electorate may have cast their votes, but only 50 per cent. of the middle and professional classes; the final result would not then conform to the straw vote taken by the poll.[1] In the same way, attempts to interpret the trend of public opinion from letters written to a particular daily newspaper or straw votes from its readers are not based on any sound statistical reasoning, since most newspapers appeal mainly to a specific section of the electorate and only those of their readers who feel especially strongly about any issue bother to write.

The problems raised by the technique of sampling are many and complicated. Usually the advice of an expert in this particular field is obtained to avoid mistakes at that stage which could nullify all the effort expended on the enquiry. Apart from the need for random selection there are no precise rules or formulae to follow, and often if there is even the slightest doubt as to the accuracy of the results of a sample enquiry, it is advisable to check the results by further sample surveys. This is particularly true where a rather small sample has been selected. To sum up, therefore, the main considerations to be borne in mind in sampling may be repeated. Firstly, the population must be homogeneous, *i.e.* only items and groups with common

[1] Despite all the adverse criticism directed against the polls on this occasion, the actual error was less then 4 per cent. of the votes cast.

characteristics relevant to the subject of the enquiry may be compared. Secondly, the sample must be both random and large enough to be fully representative of the whole group, *i.e.* large enough to contain all significant differences in the characteristic under review. Finally, in collecting the sample, the temptation to take the easiest and most accessible items must be overcome. For instance, if a poll is to be taken of student's views on the suitability of a certain hour for a lecture, it should not be conducted before the lecture starts unless all are present. The very reason for the absence of some of those not there to express their views might be the unsuitability of the hour. Thus the result of the poll would be distorted, as will the result of any sample survey in which the respondents are not selected at random. Similarly an investigator refused information at one of the households selected for an enquiry should not go next door, even if invited to do so, since the answer provided by an unwilling informant may be very different, and sometimes less prejudiced, *i.e.* more representative of the body of opinion, than those supplied by a willing neighbour who may have an axe to grind!

Some Recent Schedules

To illustrate the points and comments on questionnaires made in the foregoing chapter, four illustrations of schedules are discussed in this appendix.

1. The first is a very simple schedule used by the Board of Trade in the Census of Production. This census is held about every three years, and in the intervening years a sample enquiry is carried out. The form reproduced here is sent to all small undertakings, which for the 1958 census are defined as firm employing less than 25 persons including working proprietors. The amount of information required is very small, it consists merely of a description of the firm's main line of business and the number of its workpeople. This schedule should be compared with the last of the four in this section, which is the form sent to all establishments. Some comments on that form are made later. The reason for this very simple form is that small undertakings in many industries are very numerous yet in the aggregate produce a relatively small proportion of the industry's total output. Furthermore, such enquiries tend to be resented as wasting the proprietor's time and the non-response rate is high. Since the

BOARD OF TRADE

CENSUS OFFICE

Lime Grove, Eastcote, Ruislip, Middlesex

1st October, 1958.

Dear Sir(s), CENSUS OF PRODUCTION FOR 1958

All firms engaged in the manufacture or processing of goods, the construction or repair of buildings, etc., or any other form of productive work, including mining, will be required under the Statistics of Trade Act, 1947, to make a return to this Office in connection with the census of production for 1958.

I am anxious to avoid sending a detailed form next January to the smaller firms and I am therefore inviting you to complete the statement below if, during the period 1st October, 1957, to 30th September, 1958, the average number employed in your business (including working proprietors and clerical staff) was 24 or fewer.

If you are mainly engaged in retailing or wholesaling goods purchased from other firms and your own production is only a minor part of your business please say so at heading 1 below. If you are engaged mainly in repair work please so state at heading 1 and add the kind of goods you repair.

If you return this form satisfactorily completed without delay, you will not be further troubled in connection with the census of production for 1958. If you do not reply a detailed form will be sent to you in January for completion.

Yours faithfully,

Director of Statistics.

CONFIDENTIAL

Statement to be completed only by those firms who, during the period 1st October, 1957, to 30th September, 1958, employed not more than 24 persons.

	Codes (*leave blank*)	Machine Codes
1. Full description of work done	/ /	31.16. 9
2. Average number of persons employed in the business (including working proprietors and clerical staff).	Number	
Males	....	31.31. 5
Females	....	31.36. 5
3. If production began or ceased during the period state here :		
Date when production began	......../......../........	31.51.10
Date when production ceased	......../......../........	31.61.10

I hereby declare that the information given above is correct to the best of my knowledge and belief.

Date...1958. *Signature of person furnishing information* ....

Title of firm ....

Address of registered office (if a limited company) } ....
....

Note : No envelope or stamp is needed when returning this form ; it should be folded as explained overleaf.

FORM CP/2 (Std.)

Board of Trade can exercise its statutory powers to evoke a reply, in due course the majority of firms do make a return. But if any detail were required, as was the case in the first post-war census, the questions are carelessly answered. In all probability, the census statisticians can make as good an estimate of the output of such small firms on the basis of the numbers employed as they are likely to make on the basis of a large amount of inaccurate data supplied by unwilling informants. The reader will note the brief letter at the head of the form explaining its purpose. The implied 'threat' of further forms in the last sentence is perhaps a good method of ensuring a fairly prompt response!

2. The second schedule was used in the Family Census carried out by the Royal Commission on Population Statistics subcommittee. This form was sent to 1·7 million married women together with a brief accompanying letter explaining its purpose and, for obvious reasons, emphasising the fact that the information disclosed on the form would remain 'strictly confidential'. Two successive follow-up letters were used where the enumerator who called to collect the schedule had been told that the woman was unwilling to complete the form and another letter where the enumerator had called three times without gaining any response. This tactful approach was necessary since the census was voluntary; unlike the population census the government had no power to compel people to complete the schedule. In the event some 87 per cent of the 1·7 million women covered by the enquiry returned the form.

A feature of this form is the great care devoted to making it both simple and attractive; a commercial artist was employed to design it. The lay-out is especially good, the separate sections covering the woman herself, her children and her husband. On the whole the individual questions were very simple and incorrect answers would only arise from utter carelessness or a desire to mislead. Note the footnote in question 7a where detailed information is asked about the husband's work. This is a result of the experience of population censuses where the question on 'occupation' is a serious source of error.

3. The 'Farm Survey' form, No. B496/E1 is an interesting example of the detailed information which can be collected and compared, using interviewers who have been given explicit instructions. Section A could be answered by any farmer; but

ROYAL COMMISSION ON POPULATION

FAMILY CENSUS — Strictly Confidential

If you care to fill up this form yourself, please do so and give it completed to the Royal Commission Enumerator who will call to see you. If you prefer it, the Enumerator will be glad to fill up the form for you or help you with any difficulties you may have.

For Official Use Only

F.O..................

S.N..................

YOURSELF

Please write clearly and in full

1 Are you now Married or Widowed – *or was your last marriage ended by* Divorce? } *Please state which*

	Month	Year
2 When were you born?		

	Month	Year
3 For those who have been Married Once Only:		
(*a*) When were you married?		
(*b*) If your marriage has ended – when did it end? (*By death of your husband, or divorce – NOT separation*)		

	Month	Year
4 For those who have been Married More Than Once:		
(*a*) When were you First Married ?		
(*b*) When did your First Marriage End ?		

YOUR CHILDREN

5 (a) Number of Children Born Alive.

Beginning with your FIRST BORN child – enter, in order of birth, the date of birth of EVERY LIVE BORN CHILD you have had– *whether or not the child is still living.* Do NOT include still-births or miscarriages.

In the case of twins or triplets, use a separate line for every child born alive. Step-children or adopted children should NOT be counted.

(b) No Children
If you have NOT borne a living child, write NIL in this box.

Note: For those who have had more than 10 children there are more spaces on the back.

	Month	Year
1st child		
2nd ,,		
3rd ,,		
4th ,,		
5th ,,		
6th ,,		
7th ,,		
8th ,,		
9th ,,		
10th ,,		

6 Of your children Alive today, how many have NOT yet reached their Sixteenth birthday?
(*Only children BORNE BY YOU and under 16 – even if they are living away from you*)

For Official Use Only:

1

M 1
M 2
MM 3
MM 4
W 5
W 6

2

1 P	WM
2 E	7 F
3 OA	8 AW
4 S	9 L
5 WE	X AF

3 L.C.

4 T.C.

INTERVIEW DATE

YOUR HUSBAND

If possible, discuss this section with your Husband

7 (a) What is your Husband's Occupation?

(*If he is retired, out of work, or dead – state his former occupation*)
(*If he is temporarily in the Services – state former occupation. If no former occupation – put 'Armed Forces'*)
(*If he is a regular Sailor, Soldier or Airman – state which, and his rank*)
(*If you have been married more than once – the answer should refer to your FIRST husband*)

Note: Please describe the KIND of work your husband does in as much detail as possible. For example: if your husband is an Engineer, it will help us if you can say EXACTLY which kind he is.

(b) Is Your Husband –

1 An employer of 10 or more people?
or
2 Working for himself or employing LESS than 10 people?
or
3 Employed and earning a monthly salary?
or
4 Employed and earning a weekly or other wage?

PLEASE PUT A RING ROUND THE NUMBER WHICH APPLIES

(c) Employer's Business –

(*If your husband is NOT himself an employer or working for himself*)

FARM SURVEY

Form No. B496E./I

A. TENURE

1. Is occupier tenant .. ..
   owner .. ..

2. If tenant, name and address of owner:

....

....

3. Is farmer full time farmer ..
   part time farmer ..
   spare time farmer
   hobby farmer ..
   other type ..

Other occupation, if any:

	Yes	No
4. Does farmer occupy other land?		

Name of Holding	County	Parish

	Yes	No
5. Has farmer grazing rights over land not occupied by him?		
If so, nature of such rights:		

....

B. CONDITIONS OF FARM

1. Proportion (%) of area on which soil is ..	Heavy	Medium	Light	Peaty

2. Is farm conveniently laid out?

Yes	
Moderately ..	
No ..	

3. Proportion (%) of farm which is naturally	Good	Fair	Bad
4. Situation in regard to road . ..			
5. Situation in regard to railway ..			
6. Condition of farmhouse			
Condition of buildings			
7. Condition of farm roads			
8. Condition of fences			
9. Condition of ditches			
10. General condition of field drainage ..			
11. Condition of cottages			

	No
12. Number of cottages within farm area	
Number of cottages elsewhere ..	
13. Number of cottages let on service tenancy	

14. Is there infestation with:

	Yes	No		Yes	No
rabbits and moles			rooks and wood pigeons		
rats and mice			other birds		
			insect pests		

	Yes	No
15. Is there heavy infestation with weeds?		
If so, kinds of weeds:		

	Yes	No
16. Are there derelict fields? ..		
If so, acreage		

Code No. ....

County ....

District ....

Parish ....

Name of holding ....

....

Name of farmer ....

Address of farmer ....

....

No. and edition of 6-inch O.S. Sheet containing farmstead ....

C. WATER AND ELECTRICITY

Water supply:	Pipe	Well	Roof	Stream	None
1. To farmhouse					
2. To farm buildings					
3. To fields					

	Yes	No
4. Is there a seasonal shortage of water?		
Electricity supply:		
5. Public light		
Public power		
Private light		
Private power		
6. Is it used for household purposes?		
Is it used for farm purposes?		

D. MANAGEMENT

1. Is farmer classified as A, B or C? ..	
2. Reasons for B or C:	
old age	
lack of capital .. .	
personal failings	

If personal failings, details:

....

....

....

	Good	Fair	Poor	Bad
3. Condition of arable land ..				
4. Condition of pasture ..				

5. Use of fertilisers on:	Adequate	To some extent	Not at all
arable land			
grass land			

Field information recorded by

....

Date of recording..... ....

This primary record completed by

....

Date ....

clearly question 2 in Section B must be capable of a standardised answer, *i.e.* all interviewers must have a clear ruling as to the difference between 'Yes' and 'Moderately'. Similarly, the answers to Section D would also require care to ensure that the grades A, B and C were uniformly applied, not merely as between different farmers by the same interviewer, but more especially by different interviewers. This is a highly relevant consideration since the staff of field workers, both paid and voluntary, was drawn from the staffs of the County Agricultural Committees, and they had no statistical expertise. According to the National Farm Survey report, the usual procedure was for the recorder to carry out a general tour of inspection and answer the qualitative type of question, *e.g.* condition of farmhouse, while the farmer himself provided the quantitative information, *e.g.* number of cottages. The report admits the lack of standardisation of qualitative data, due to inexperienced staff. For reasons given above the classification is seriously affected by the differences in local standards and experience, and comparisons on a national scale are for this reason unsatisfactory.[1]

4. The final schedule is to be used for all firms in the census relating to the year 1958 which will be carried out during 1959. Note that all firms receiving it are expected to complete the first set of questions. The small firms who did not return the earlier schedule reproduced on p. 27 will now have to complete the first two boxes. The remainder of the first page of the schedule relates only to firms with more than one establishment. The reverse side of the schedule depicted on p.32 contains twenty-eight questions classified under 8 headings. All the questions asked should be fairly easy to complete since they entail no more than extracting the relevant figures from the company's financial records. This form is used for the larger establishments which it may be assumed possess an accounts department which can be relied upon to furnish the information required without much trouble and with reasonable accuracy. The spirit of co-operation does seem to vary, however, from firm to firm!

If there is a serious criticism of this form it concerns the footnote asterisk on the front page instructing the reader to refer to note 15 as well as note (i) on the second page asking that the

[1] National Farm Survey of England and Wales. H.M.S.O., 1946, pp. 3-6.

CENSUS OF PRODUCTION FOR 1958

MOTOR VEHICLES AND CYCLES (MANUFACTURING)

INFORMATION COPY

NOT TO BE

COMPLETED

CONFIDENTIAL

The above numbers should be quoted in any correspondence

Board of Trade,
Census Office,
Lime Grove,
Eastcote, Ruislip,
Middlesex

Pinner 9800 Ext. 28

If the name or address shown above is incorrect in any respect, please correct it.

DETAILS OF BUSINESS

To be completed by ALL FIRMS

1. State the name of the firm carrying on business at the above establishment ....
2. State the principal trade or business carried on at the establishment .. ..
3. If your firm is a limited company, state the address of its registered office. If it is not, state the full name(s) of proprietor(s) .. ..
4. If you have ceased to carry on business at the above establishment, state here the date when you ceased . ...
5. State the year of return (*see note* 8). Twelve months ended ....195.. ....
6. Has any capital expenditure been incurred at any establishment, under the same control as that covered by this return, which had not begun production by 31st December, 1958, and for which, therefore, a Census of Production return has not been completed? If so, state its address here. ....

SMALL FIRMS

To be completed by firms that employed FEWER THAN TWENTY-FIVE PERSONS *on the average during* 1958.

Nature of work done ....

Average* number of persons employed during 1958 inside and outside the factory or workshop:

Males .... ( ) Females .... ( )

Include working proprietors and clerical staff. Exclude outworkers.

All returns must be signed at the end *but the details in the rest of the return should be completed only by firms that employed* TWENTY-FIVE OR MORE PERSONS *on the average during* 1958.

ESTABLISHMENT TABLE

To be completed if this return covers more than one establishment (*see notes* 7(*a*) *and* 9)

For official use	Address	Nature of work carried on	Average* number of operatives employed
			No.
		Total	

* *See note* 15

(i) Please read the notes sent with the form before completing the return
(ii) All figures should relate to the year of return
(iii) State values to the nearest £
(iv) Do not leave blanks : where none state " none "

CONFIDENTIAL

	Machine codes
I WORKING PROPRIETORS (*see notes* 10 *and* 11)	
1. **Number : Male**	
2. **Female**	
II EMPLOYMENT (*see notes* 12-15)	
A. Number of persons employed in the pay-week ended on or about 25th October, 1958:	
(i) Operatives :	
3. Male	
4. Female	
(ii) Administrative, technical and clerical employees :	
5. Male	
6. Female	
B. Average* number of persons on the pay-roll:	
7. Operatives	
8. Administrative, technical and clerical employees	
III WAGES & SALARIES (*see notes* 16 *and* 17)	
A. Paid during the year to:	
9. Operatives £	
10. Administrative, technical and clerical employees £	
B. Salaries, etc. paid to administrative, technical and clerical employees in October, 1958:	
(i) Staff paid monthly : amount paid for October, 1958.	
11. Male £	
12. Female £	
(ii) Staff paid weekly : amount paid for week ended on or about 25th October, 1958.	
13. Male £	
14. Female £	
IV MATERIALS AND FUEL PURCHASED (*see notes* 18-20)	
15. Total cost £	

	Machine codes
V WORK GIVEN OUT (*see notes* 21-23)	
16. Total amount paid £	
VI TRANSPORT PAYMENTS (*see notes* 24-26)	
17. Total amount paid (or credited) £	
VII STOCKS (*see notes* 27-33)	
Materials and Fuel:	
18. At beginning of year £	
19. At end of year £	
Work in progress:	
20. At beginning of year £	
21. At end of year £	
Products on hand for sale:	
22. At beginning of year £	
23. At end of year £	
VIII CAPITAL EXPENDITURE (*see notes* 34-40)	
Plant, Machinery and Vehicles	
Cost of items acquired:	
24. Plant and machinery £	
25. Vehicles £	
Proceeds of items disposed of:	
26. Plant and machinery £	
27. Vehicles £	
New Building Work	
28. Cost of new building or other constructional work of a capital nature charged to capital account during the year £	

**See note* 15

accompanying notes should be read. The notes in question cover four foolscap pages of print and while it is highly desirable on grounds of statistical accuracy that all informants should interpret questions in the correct manner, it is the height of optimism to assume that the majority of informants study these notes. It would have been better if such questions as need interpretation had been supplemented by a small footnote in the same 'box' explaining what is required. The present notes are likely to be 'overlooked'! The census statisticians probably console themselves with the thought that since most industries are dominated by about half-a-dozen giant concerns whose staffs are more accustomed to such forms, the data collected from them will ensure a reasonably accurate picture of the industry. In any case the numerous errors will probably tend to cancel each other out. It would be unduly optimistic, however, to pretend that the smaller concerns give these forms the care and attention they need.

All who seek information from others should bear in mind the guiding principle that it is the informant who is doing you the favour of giving information; it is wise policy to make the task easy for him.

CHAPTER IV

TABULATION

The Purpose of Tabulation

In no investigation of any size is the volume of collected data or material so small that it may be rapidly or easily assimilated by a perusal of the completed forms. At best only the haziest impressions may be gathered of the ultimate results, and those impressions may well be the reverse of the truth for it is usually the unusual or freak cases that stick in the memory to the exclusion of the many more 'ordinary' replies. The statistician's first task is to reduce and simplify the detail into such a form that the salient features may be brought out, while still facilitating the interpretation of the assembled data. This procedure is known as classifying and tabulating the data; *i.e.* extracting from the individual questionnaires the answers to each question and entering the replies on separate summary sheets. These totals are then transferred to the relevant columns of prepared tables.

Before the summarising is commenced, the questionnaires should have been checked on receipt to ensure that they have been completed reasonably correctly.[1] Inevitably there will be uncompleted forms, and these it may be possible to complete by further enquiry; in other cases the replies will be useless, as they are either irrelevant or patently false. It requires little imagination on the reader's part to visualise the task involved in sorting several thousands of forms and tabulating their contents without mechanical assistance. The results of many large-scale enquiries would be available only long after the field work had been completed. Fortunately the introduction of mechanical punching and sorting machines has facilitated the task of the statistician. All that is now necessary is for the information on the questionnaires to be transferred to specially designed punched cards, the machines then sort the cards, tabulate and compute the totals at great speeds. There is, of course, the risk that the cards may be incorrectly punched by the operator, but this problem can be overcome with adequate supervision. The cost of such methods

[1] The person editing the returned forms cannot know if they are *correctly* completed, otherwise there would be no need for the enquiry. 'Reasonably correct' in this context means that there are no obvious mistakes in, or contradictions between, the various answers.

is considerable and careful thought must be given to the information to be entered on the cards to ensure that the maximum information will be given out by the machines in the minimum space of time.[1] According to an article in the Journal of the Royal Statistical Society,[2] the representative of a machine accounting company was able to devise a procedure whereby the first results of the census of the population of Cyprus were available within a few weeks, instead of several months, as is usually the case.

The Basis of Classification

Before the actual tabulation can be undertaken there is an intermediate stage, generally described as classification. The point has already been made that statistics is concerned with aggregates, the individual members of which are homogeneous. That is, all the items comprising the aggregate or what the statistician calls the 'population' are of one type, *i.e.*, they possess a characteristic in common. For example, retail businesses may be classified according to turnover, schoolboys according to their heights, shares quoted on the Stock Exchange according to the dividend paid. The 'characteristics' in these cases are: turnover, height and dividends respectively, *i.e.* these constitute the link or basis of comparison between all the items within each group. These groups of individual items are usually termed statistical *series* or *distributions*. A more precise way of defining a series or distribution is that it comprises a group of items which are related one to the other by the possession of some common characteristic.

The term 'series' is usually restricted to data which have been collected over time. The figures of the annual turnover of a firm for the past decade would be described as a *time* series. Data which relate to any characteristic other than time may be classified as spatial or attributive distributions. Whereas the time series indicated the turnover of a group of departmental stores over a period of several years, a *spatial* distribution may be one which classifies the turnover of any period according to the location of the sales. Thus, the turnover may be classified according to the departments of the stores, or on the basis of comparing

[1] J.R.S.S., 1946, Part III, p. 284. An article by O. Kempthorne illustrates the use of mechanical methods in some detail, with special reference to the National Farm Survey data. An appendix on mechanised punched card systems is included at the end of this text.

[2] J.R.S.S., 1947, Part II, p. 138. An experiment in census tabulation. – D. A. Percival.

the annual turnover of the various stores in the different towns. In brief, spatial distributions are concerned with location. The term '*attributive*' covers all distributions other than time series and spatial distributions. An attribute is simply a characteristic; data are classified according to their attributes. As already explained (p. 11) these fall into two main types, those which are capable of numerical expression, and those which are not. Thus, in the example of schoolboys classified according to height, the characteristic can be expressed numerically, *i.e.* so many inches. But if the boys are being graded by the school doctor on the basis of their general health and physique, they could only be classified somewhat as follows: excellent, good, fair, and poor. The first example of an attribute, which can be expressed in quantitative terms, *i.e.* height in inches, is termed a *variable;* the second type of classification is based on the *attribute* itself, *i.e.* a quality incapable of being measured in numerical terms, such as 'good health', and is given that name. These points have been repeated because they are relevant not merely to classification; they also affect the type of tabulation used and the form of diagrammatic representation of the data.

The Construction of Tables

The purpose of tabulation is primarily to condense and thereby facilitate comparison of the data. The form of the tables employed will vary according to the nature of the data and the requirements of the survey. In consequence, it is not possible to lay down hard and fast rules which may be applied in all cases. It may come as a surprise to the reader to learn that the tables are usually drawn up before the enquiry is actually started. More precisely, the frame of the tables is drawn up and this has two advantages. First, it enables the survey team to visualise the sort of data they want and are going to get and second, it sometimes draws attention to other information which would be of interest and provision for such questions can then be made on the questionnaire. As with so many matters, common sense dictates certain guides which should be borne in mind in the construction of any statistical table if it is to serve the purpose of revealing the basic structure of the data.

In no case should the table be overloaded with detail. A closely-printed and concentrated mass of figures may appear

most impressive to the casual observer, but merely compels the reader to do what ought to have been done by the compiler of the table at the outset: namely, to reduce the mass into several sub-tables, each bringing out a separate aspect of the data. The purpose of the table should be immediately apparent, *i.e.*, it should have a clear and concise title, although clarity and precision should not be sacrificed for brevity. Occasionally tables are encountered where the main title is amplified by a series of footnotes; wherever possible, this practice should be avoided. Where the individual figures are large, the table gains in clarity far more than it loses by eliminating the '000's' or even '00,000's', *i.e.*, the final digits. This is especially true of summary tables which are often inserted in the text in the body of the main report or its conclusions; individuals seeking detailed figures can be referred back to the full tables which are best put into an appendix separate from the main report.

It is highly desirable in any report presenting data collected for that enquiry to precede the information presented in tabular form by a short summary of the methods of collection employed, in order that the reader may obtain some idea as to the probable reliability of the results given in the tables. If secondary data from other published sources are given, say for purposes of comparison, then a footnote should be appended indicating the source of that particular section of the table. Especial care should be taken to leave the reader in no doubt as to the unit of measurement: whether it be £ sterling or £ Australian, long or short tons, ton-miles of passenger trains, or goods trains, etc. If any heading is at all liable to misinterpretation, a clear definition should be provided as to what information is included under that head. Thus, the statistics published by the Home Office of 'persons proceeded against for drunkenness' do not include those persons charged with 'driving under the influence of drink'; these are incorporated with offences against the Highway Act.

Simple Tabulation

To illustrate the normal procedure in tabulation, the data given in Table 1 relating to the individual outputs of 180 workers producing a certain manufactured article are set out with the smallest output at the beginning of the group, and the largest at

the end, *i.e.*, in order of magnitude. Such an arrangement of the data is known as an 'array'. Inspection of the table reveals that the minimum and maximum outputs are 501 and 579 respectively. The difference between these two quantities is described as the *range*. Apart from the range, it is impossible without further careful study to extract any exact information of any value from the table. By breaking down the data into the form of

TABLE 1

GREAT PRODUCERS LTD.

INDIVIDUAL OUTPUTS OF 180 FEMALE OPERATIVES IN PLANT 1, IN THE WEEK ENDING 8TH NOVEMBER, 1958

501	520	534	540	547	555
503	522	535	542	547	557
503	522	535	542	547	557
504	523	535	542	547	557
506	523	535	542	547	559
507	524	536	542	548	559
507	525	536	542	548	559
509	525	537	543	548	559
510	526	537	543	548	559
511	526	537	543	549	561
511	527	537	543	549	561
513	527	537	544	549	561
515	527	537	544	549	563
515	528	538	544	550	563
515	528	538	544	550	563
515	528	538	544	550	564
515	528	538	544	550	565
515	528	538	545	551	565
517	528	539	545	551	565
517	530	539	545	551	567
518	530	539	545	551	567
518	532	539	546	552	567
519	532	539	546	552	569
519	532	539	546	552	569
519	532	539	546	553	569
519	532	539	546	553	572
520	532	540	546	553	574
520	534	540	547	553	575
520	534	540	547	555	577
520	534	540	547	555	579

Table 2 below, however, certain features of the data become apparent. Thus, by setting the number of workers producing each individual quantity against that figure, a more intelligible picture is provided. Even a superficial scrutiny of Table 2 reveals that the ouputs from 537 to 547 inclusive occur most frequently.

Such a table is known as a *frequency distribution*. It is so described because it indicates the frequency or number of times each individual output figure occurs. More precisely, it tabulates the frequency of occurrence of the different possible values of

TABLE 2

FREQUENCY DISTRIBUTION OF OUTPUTS DETAILED IN TABLE 1

Output	Frequency	Output	Frequency	Output	Frequency
501	1	527	3	550	4
503	2	528	6	551	4
504	1	530	2	552	3
506	1	532	6	553	4
507	2	534	4	555	3
509	1	535	4	557	3
510	1	536	2	559	5
511	2	537	6	561	3
513	1	538	5	563	3
515	6	539	8	564	1
517	2	540	5	565	3
518	2	542	6	567	3
519	4	543	4	569	3
520	5	544	6	572	1
522	2	545	4	574	1
523	2	546	6	575	1
524	1	547	8	577	1
525	2	548	4	579	1
526	2	549	4		

TABLE 3

GROUPED FREQUENCY DISTRIBUTION. DATA FROM TABLE 2

Output	No. of Operatives
(Units per operative)	
500— 9	8
510—19	18
520—29	23
530—39	37
540—49	47
550—59	26
560—69	16
570—79	5
	180

any given variable. Nevertheless, even after this simplification, since there are still too many figures to assimilate, the conventional procedure is to construct a *frequency table* as in Table 3.

The data in this form are sometimes described as a '*grouped*'

frequency distribution.[1] Instead of the 'frequencies' of each output being shown separately, the range (difference between maximum and minimum outputs) is sub-divided into smaller groups, usually termed 'classes'. In this example each class comprises ten units of output. Thus the first class, 500-509, covers all ten values inclusive. In this class eight operatives had outputs of 500 units or above, but none more than 509. In the fifth class, 540-549, there were 47 operatives, none of whose individual outputs was below 540 or exceeded 549. The reader can verify the figures in Table 3 by reference to the previous table.

By grouping the data into the form of such a frequency table, the basic structure of the information is prominently revealed. The main body of operatives have outputs falling within the middle classes, and only a few operatives come within the classes at either extreme. In passing, it can be stated that the frequency table or 'grouped' frequency distribution is the most common form of presentation of numerical data, and, as will be seen later, is the basis of most statistical analysis.

The same data can be presented in cumulative form, *i.e.*, 86 operatives each produced less than 540 units per week, 133 operatives less than 550; and so on up to the last stage, when 180 operatives each produced less than 580 units. Note that the 'cumulation' may be upward or downward; the upper half of Table 4 is read as '133 operatives produced less than 550 units per week . . . ' etc.; the lower half as '94 operatives produced 540 or more units per week'.

TABLE 4

DATA FROM TABLE 3 PRESENTED IN CUMULATIVE FORM

Output (Units per operative)	500–9	510–9	520–9	530–9	540–9	550–9	560–9	570–9
No. of operatives	8	26	49	86	133	159	175	180
	180	172	154	131	94	47	21	5

This table yields the data to answer such questions as 'What percentage or proportion of the workers produce 550 or more units per week?' In this case the answer would be 26 per cent. (*i.e.*, $\frac{47}{180} \times 100$).

[1] The term generally used is 'frequency distribution.' *i.e.* the same term is usually used whether the data are grouped or not.

The Selection of 'Classes'

The preparation of grouped frequency distributions from the raw data may give rise to difficulties. The greatest of these is deciding the number of *classes* into which the series may be divided: *e.g.* in the above table there are eight classes, 500-9, 510-9, and so on up to 570-9. This problem is directly linked with the second: what is the size, or, more precisely, the range, of each class to be, *i.e.* what is the *class-interval?* Thus, in the above example, the class-interval is ten units. There are no hard and fast rules on these points, but generally it is desirable that the number of classes should not exceed 15 or 20, depending on the range of the variable and total frequencies in the distribution, otherwise the purpose of the table, the reduction of the data to manageable size, may be defeated. As to the size of the class-interval, this will depend primarily on the number of classes and the distribution of the frequencies. Sometimes the class-interval is easily determined, *e.g.* if families are being classified according to the number of children in each, then the class interval is clearly one child. The frequency table heading would read:

Number of children per family	No. of Families

There are no set rules which if followed will ensure good tabulation in all cases. The student should at all times bear in mind the two main needs of tabulation work. First the table must be comprehensible in that it reduces the data to manageable proportions, and second, the content of the table is clearly yet simply defined in the title and column headings. In brief, Table 1 on p. 39 may have its uses for detailed analysis of the factory's production performance; but Table 3 is far more comprehensible and its content easier to grasp. Most tables should resemble Table 3 rather than Table 1.

Where the individual values to be classified tend to group themselves around particular values, care should be taken when deciding upon the class-interval that such values of the independent variable coincide as far as possible with the mid-points of the classes. For example, if in Table 1 a large majority of the outputs ended with the digit 5, *e.g.* 545, 555, etc., the classification used in Table 3 is excellent. If, however, there were many outputs ending in 9, 0 or 1, then it would be preferable to draw up classes as follows: 505-514, 515-524, 525-534 and so on.

This is because it is customary to regard the middle values of classes as the 'average' value of the items in that class. If, as will be shown in Chapter VI, calculations are to be carried out on the data, any inaccuracy or bias in the grouping within the class intervals may distort the final results. To illustrate the effect of using the same-sized class interval (10 units) with different limits, Table 5 has been drawn up showing the data from Table 1 classified in two ways. The differences can be observed by comparing the resultant distributions.

TABLE 5

EFFECTS OF DIFFERENT CLASSIFICATION ON DISTRIBUTION OF DATA FROM TABLE 1

First Grouping	Frequencies	Frequencies	Alternate Grouping
—	—	4	Under 505
500— 9	8	8	505—514
510—19	18	24	515— 24
520—29	23	25	525— 34
530—39	37	46	535— 44
540—49	47	41	545— 54
550—59	26	18	555— 64
560—69	16	11	565— 74
570—79	5	3	575— 84
	180	180	

Further reference will be made to this matter when the various averages are discussed, but keeping the above principles in mind, the numerous types of tabulation and classification can be examined. Table 3 above, *i.e.*, the frequency table, is the basic and most simple form of presenting data, and even the most complex tabulation comprises little more than a number of such tables brought together under one head. The more information which it is sought to bring into one table, the more important it becomes to ensure that the table remains intelligible and easily readable.

Further Examples of Tabulation

A number of tables taken from official publications are reproduced in the following sections, together with comments upon their salient features.

Table 6 is an illustration of double tabulation describing the distribution of personal incomes in Great Britain in the fiscal year ended 31st March 1956. The class interval, it will be noted, is not even, but is adequate to illustrate the distribution of incomes in this country. When the same data for the fiscal year 1952-3 were published in the 97th Report, the class-interval was as follows: £250-500, 500-750, 750-1,000 etc. This illustrated a serious error of classification, for if such a grouping is used, it creates doubt as to the correct treatment of border-line incomes, *e.g.* £500, £750 and £1,000. Are they in the class with those figures as the upper limit, or in the next class where the figure is the lower limit? The lower and upper limits of two successive classes may appear to be the same, *e.g.* £250 and under 500, £500 and under 750, etc. In such cases, however, no uncertainty arises as to the treatment of income of £500, 750, etc. In Table 6 the Revenue statisticians have avoided the earlier mistake by defining precisely the actual limits of each class.

TABLE 6

DISTRIBUTION OF PERSONAL INCOMES BEFORE TAX 1955-56

Range of Total Income	No. of Incomes	Total Income before Tax
£	000's	£ million
180— 249	2,075	443
250— 499	7,500	2,837
500— 799	7,300	4,607
800— 999	1,800	1,587
1,000—1,499	925	1,094
1,500—1,999	260	446
2,000—2,999	175	422
3,000—4,999	104	393
5,000 and over	61	545
	20,200	12,374

Source: Based on Table 56 in 100th Report of the Commissioners of H.M. Inland Revenue, Cmnd. 341, H.M.S.O. 1958.

When interpreting the table it should be borne in mind that these are, in fact, *declared* personal incomes. To the extent that a return of income is not made by various individuals who wish to avoid their obligations to the Revenue, or where the income declared is incorrectly stated, the table fails to reflect the true state of affairs. Since the number of individuals in these two

categories are relatively small in relation to the tax-paying population the error is probably not significant.

It is interesting to note that less than eight people in every hundred earned over £1,000 a year in 1956 and that nearly half the personal incomes were below the £500 a year level. As an exercise the student reader could calculate the proportion of incomes in each income group and their share of the total incomes returned. For example, less than 2 per cent (340 out of 20,200) receive just under 11 per cent of total incomes (£1,360 out of £12,374). The results are interesting.

Table 7 illustrates how neatly a great deal of data can be compressed and presented in a single table with a two-way classification, *i.e.* number of incomes vertically and size of household incomes horizontally. It would be instructive to write a prose passage containing all these figures and to compare it with the table. The advantages of the tabular form would be immediately apparent.

Table 7 explains in part why, despite the fact as revealed in Table 6 that about 50 per cent of taxpayers receive less than £500 p.a., the standard of life among many such working class families is high. It is clear that in more than half the households covered by this official enquiry, there are at least two incomes coming into the home. Since this sample of 12,911 households constitutes a representative cross-section of the community, it

TABLE 7

ANALYSIS OF HOUSEHOLDS IN DIFFERENT INCOME GROUPS BY NUMBER OF INCOME RECIPIENTS

Number of income recipients in household	£20 or over	£14 but under £20	£10 but under £14	£8 but under £10	£6 but under £8	£3 but under £6	Under £3	TOTAL
One ..	251	603	1,445	1,329	1,034	688	678	6,028
Two ..	414	1,173	1,593	604	357	561	69	4,771
Three ..	397	610	351	93	41	30	—	1,522
Four ..	250	159	32	5	5	—	—	451
Five ..	75	30	3	—	—	—	—	108
Six or over	27	3	1	—	—	—	—	31
TOTAL ..	1,414	2,578	3,425	2,031	1,437	1,279	747	12,911

Source: Table 5 in Report of an Enquiry into Household Expenditure in 1953-54. H.M.S.O. 1957.

may be inferred that in about one-half of all households in the United Kingdom there are at least two incomes. For this reason it is not possible to generalise about living standards in the United Kingdom solely by reference to wage levels and the cost of living.

Table 8 reveals yet another aspect of the community's economic behaviour, for while Table 6 revealed the distribution of incomes and Table 7 showed what the 13 million households in the United Kingdom made up of the 20·2 million income recipients had to spend each week, Table 8 shows how they spent their money. The reader should note that these figures cannot be directly related to the totals in Table 6 because the latter are given before tax, and those in Table 8 after tax and personal savings.

TABLE 8

CONSUMERS' EXPENDITURE UNITED KINGDOM 1953-1957

(£ million at 1948 market prices)

Outlay	1953	1954	1955	1956	1957
Food	2,492	2,555	2,592	2,630	2,669
Alcoholic drink	838	849	882	904	921
Tobacco	794	811	831	842	865
Housing, fuel and light ..	1,145	1,193	1,186	1,206	1,207
Durable household goods ..	624	714	762	736	801
Clothing and footwear ..	912	968	1,030	1,061	1,074
Private motoring and cycling..	272	328	416	364	369
Other goods	638	685	712	717	744
Other services	1,354	1,369	1,367	1,360	1,365
TOTAL ..	9,069	9,472	9,778	9,820	10,015

Source: Economic Survey 1958. Cmnd. 394.

There are two points of interest for the statistician. The first is the technique of converting all the money totals into their 1948 equivalent; in other words for any category of expenditure, *e.g.* food, changes in the annual totals reflect *real* or quantity changes in consumption. The second point concerns the source of these figures and their reliability. Some of these figures are certainly more reliable than others. For example, the Customs and Excise figures of the sales of dutiable tobacco and alcohol provide a firm basis for the figures of total consumer expenditure shown in Table 8. The figure for 'private motoring and cycling' is probably subject to a rather larger margin of error since it must

be compiled from estimates derived from various sources, *e.g.* petroleum duty paid, garage turnover as shown in the Census of Production, licences for new cars, etc. The aura of precision lent to such published data by giving the annual figures to the nearest £ million is rather misleading. The reliability of these estimates varies considerably as between categories and even as between years for any given category, *e.g.* if census data are available for one year, it helps to give a better estimate. Either such tables should be accompanied by a reference to the probable margin of error or the figures should be approximated to the nearest £10 million, or even £100 million.

Table 9 presents part of the information derived from the 1954 Census of Production. The selected table relates to a particular industry; the Census covers all manufacturing industry and similar data are published for each industry. It will be noted that this table relates to 'larger establishments only', which for purposes of the 1954 Census were defined as firms employing more than 10 workpeople. The table is virtually self-explanatory

TABLE 9

ANALYSIS BY SIZE OF RADIO AND TELECOMMUNICATIONS INDUSTRY 1954
LARGER ESTABLISHMENTS ONLY. PRIVATE FIRMS IN THE UNITED KINGDOM

Average No. Employed	Establishments	Net Output	Employees	
			Operatives	Others
	Number	£'000	Number	Number
11— 24	92	1,321	1,433	411
25— 49	104	2,531	2,991	758
50— 99	81	4,233	4,491	1,361
100— 199	82	8,045	9,014	2,820
200— 299	33	5,535	6,083	2,067
300— 399	27	6,114	6,442	2,409
400— 499	15	4,522	4,757	1,950
500— 749	32	14,500	14,835	4,885
750— 999	17	11,316	11,449	3,260
1,000—1,499	14	10,930	12,842	4,134
1,500—1,999	9	10,424	12,192	3,682
2,000—2,499	6	9,039	10,070	3,247
2,500—2,999	3	7,202	6,373	1,884
3,000—4,999	4	11,612	10,452	5,894
5,000 and over	7	31,647	32,294	15,511
TOTAL PRIVATE FIRMS	526	138,972	145,718	54,273

Source: Report on the Census of Production for 1954. Volume 4, Industry M (H.M.S.O. 1958).

although the term 'net output' is defined in a special sense for this Census. It means the value of the firm's contribution to the national product. More simply, it is the difference between the money value of the firm's total production less the costs of materials and fuel used, the former costs covering all manufacturing charges from consumable tools and plant repairs to packing materials. Note that labour costs are not deducted from the value of the gross output to arrive at the net output.

The data as presented in Table 9 need to be further analysed before the full value of the information contained therein becomes apparent. The student will note that the first two columns form an extended frequency distribution in which the independent variable (or characteristic) is the number of employees. The reader might try to compress the table by amalgamating certain classes to bring out the most important groups in terms of output. The relationship between 'operatives' and 'other employees' for differing sizes of firm also varies quite significantly and calculation of a few percentages may suggest a number of ideas to the student. On the other hand, it is common knowledge that figures of net output are liable to substantial error. The Board of Trade admits that it regularly encounters resistance on the part of firms asked to co-operate in the census and it is doubtful if the forms are completed with as much care as the statistician would wish. Such considerations demand extra special care in the employment of such statistics as these.

One of the most important sources of statistical information is the decennial census of population. This is supplemented by a system of registration through local offices of the General Register Office. Table 10 shows the regional distribution of the population of England & Wales on the occasion of two censuses, those of 1911 and 1931. Comparable figures for 1955 are derived from the estimates of the population at 30th June prepared each year by the Registrar-General. The various regions are defined in detail in the official reports and for purposes of this table the G.R.O. has had to compile the figures for the two earlier dates specially, since these 'standard' regions as they are known have been re-defined since 1939.

From the economic or sociological point of view the table is interesting since it indicates the main movements over some four decades of the population. Changes in the percentage

TABLE 10

REGIONAL DISTRIBUTION OF THE POPULATION OF ENGLAND AND WALES, 1911, 1931 and 1955.

Standard Region	1911		1931		1955	
	Number 000's	Per cent.	Number 000's	Per cent.	Number 000's	Per cent.
Northern	2,816	*7·8*	3·041	*7·6*	3,160	*7·1*
East and West Ridings	3,560	*9·9*	3,920	*9·8*	4,106	*9·2*
North Western ..	5,780	*16·0*	6,196	*15·5*	6,449	*14·5*
North Midland ..	2,640	*7·3*	2,946	*7·4*	3,456	*7·8*
Midland	3,275	*9·0*	3,743	*9·4*	4,512	*10·2*
Eastern	2,099	*5·8*	2,424	*6·1*	3,316	*7·5*
London and S. Eastern	9,107	*25·2*	10,339	*25·9*	10,962	*24·7*
Southern	1,864	*5·1*	2,135	*5·3*	2,804	*6·3*
South Western ..	2,509	*7·1*	2,615	*6·5*	3,073	*6·9*
Wales	2,421	*6·8*	2,593	*6·5*	2,603	*5·9*
England and Wales	36,071	*100·0*	39,952	*100·0*	44,441	*100·0*

Source: Annual Abstract of Statistics No. 93, Based on Table 11.

figures for each area indicate the extent to which the relative importance (as measured solely by the population size) of these regions has fluctuated. One further point on tabulation should be noted. By printing the columns of percentage figures in italics, the readability of the table is improved, not least because comparison between related columns, *e.g.* the three percentages for any area is made much easier.

Table 11 is interesting not merely as a table which presents much data of considerable interest, but it also illustrates some of the weaknesses of tabulation when a lot of information has to be compressed into a single table. Only a part of the full table published in the 100th report of the Commissioners of Inland Revenue is given; only that part relating to estates of deceased persons valued between £50,000 and £500,000. The grouping of the estates by value does not coincide with the classification used for rates of duty, hence in the second column headed 'Rate of Estate Duty payable' two figures are given in two classes. The headings of the various columns indicating the types of asset held have to be amplified by a series of footnotes, since a number of the columns given in the published report has been reduced by combining them. As a general rule in tabular work, footnotes are to be avoided.

TABLE 11

Average Distribution of Gross Assets of Estates in Great Britain in 1956-7 Liable to Estate Duty at Rates Between 35 and 65 per cent.

Range of net capital value of estate[1] £000's	Rate of Estate Duty payable Percentage	Distribution of Gross Assets (in percentages)						Total[4] Gross Capital Value
		Cash and Liquid Assets[2]	Govt. and Municipal Securities	Trade Assets[3]	Shares in other companies	House Property Land and business premises	Other assets	
Over 50 to 60	35	18·4	16·3	10·7	39·0	10·7	4·9	100·0
Over 60 to 80	40–45	19·2	15·4	9·5	39·8	10·8	5·3	100·0
Over 80 to 100	45	16·4	16·2	9·9	41·4	10·0	6·1	100·0
Over 100 to 200	50–55	14·8	16·6	12·1	41·9	8·8	5·8	100·0
Over 200 to 300	60	17·2	14·4	10·7	39·6	10·0	8·1	100·0
Over 300 to 500	65	11·1	16·2	5·4	45·8	13·5	8·0	100·0
All estates[5]	22	27·4	17·9	7·6	24·2	17·7	5·1	100·0

Source: 100th Annual Report of H.M. Commissioners of Inland Revenue Cmnd. 341. Based on Table 203.

1 The aggregate value of all assets for probate less permissible deductions, *e.g.* debts, burial expenses.
2 Includes insurance policies, shares and deposits in building societies, mortgages, bonds, etc.
3 Including shares in private British companies.
4 The percentages may not agree exactly with totals of sub-groups due to rounding of figures.
5 Including estates not shown in this table.

TABLE 12

OUTPUT, ATTENDANCE AND PRODUCTIVITY OF ALL WAGE-EARNERS IN DEEP-MINED COAL PRODUCTION

Period	Total output of saleable coal	Saturday output	Average number of wage earners on colliery books		Manshifts worked		Average number of shifts worked per week per wage earner on colliery books (adjusted)	Absence percentage				Output per manshift worked (Adjusted)
			As recorded	Adjusted	As recorded	Adjusted		Voluntary	Involuntary	Total: As recorded	Total: Adjusted	
	Thousand tons		Thousands									Tons
1951	211,883	11,825	698·6	694·6	174,650	172,650	4·78	5·62	6·53	12·15	12·21	1·224
1952	214,324	10,437	715·8	711·8	178,322	177,326	4·76	5·87	6·16	12·03	12·09	1·199
1953	212,503	11,654	712·9	712·9	173,056	173,056	4·67	4·26	8·15	12·41	12·41	1·224
1954	213,994	11,856	707·2	707·2	173,327	173,327	4·71	4·01	8·20	12·21	12·21	1·232
1955	210,260	11,812	704·1	704·1	171,261	171,261	4·68	4·12	8·42	12·54	12·54	1·227
1956	209,925	11,222	703·4	703·4	170,167	170,167	4·65	4·25	8·67	12·92	12·92	1·232
1957	210,059	10,902	710·1	710·1	170,334	170,334	4·61	6·12	7·69	13·81	13·81	1·231
1958	201,451	3,611	698·8	698·8	158,992	158,992	4·38	6·41	7·73	14·14	14·14	1·264

Source: Ministry of Power.

Table 12 is an interesting example of the need for ensuring that definitions remain unchanged in any period over which comparisons have to be made. In the table every single series has been affected by some change in definition or policy that makes comparisons or interpretation of changes a task needing great care. In such cases footnotes are essential, but as the reader will see, this table requires very careful study to avoid errors in interpreting the data. The Ministry of Power have attempted to eliminate some of the breaks in continuity by providing adjusted series but they say that owing to the lack of precise information relating to the effect on previous years of changes of definition the adjusted figures should be accepted with reserve.

As a practical example of the problems that can arise consider the figures in a little more detail:

Col. 1. Saleable Output: 1951-1955 refers to calendar years subsequently to 52 week years. 1951-2 figures refer to all coal mines in the country while 1953 and subsequent figures relate to N.C.B. mines only.

Col. 2. Saturday Output: As for Col. 1, and in addition Saturday working was suspended in 1958.

Col. 3. Average Wage Earners on Books: From 1954 a new definition of what colliery activity consisted, and a new method of counting men on books was introduced.

Cols. 5 and 7. Manshifts Worked: As for Col. 3.

Cols. 8-11. Absence Percentage: As for Col. 3 and in addition as medical certificates were no longer required for some absence after 1st June, 1957, the subsequent voluntary absence includes some absence that was previously classed as involuntary.

Col. 12. Output per manshift worked: The unadjusted figures are affected as in Col. 3 above.

At first sight Table 13 is extremely complex and to that extent deserving of criticism. It is taken from the very informative report of the Commissioners of H.M. Inland Revenue and is designed to provide the interested reader with a great deal of information. In fact, most students of taxation would be interested only in a part of this table. Table 13 is in effect four tables in one. The first relates to the number of surtax payers in each class of income. These classes, it will be noticed, are uneven in size, the class-limits being determined by the levels at which

TABLE 13

SURTAX 1955-56: ANALYSIS OF SURTAX CASES BY REFERENCE TO RATIO OF EARNED INCOME TO TOTAL INCOME IN RANGES OF TOTAL INCOME

(Assessments made up to 30th June 1957)

Range of total income			Numbers where the percentage ratio of earned income to total income is:										Total
			0—10	10—20	20—30	30—40	40—50	50—60	60—70	70—80	80—90	90—100	
Exceeding £	Not Exceeding £												
2,000	2,500		10,606	1,661	1,640	171,7	21,18	2,387	3,409	5,210	10,669	47,653	87,070
		Per cent.	*12·2*	*1·9*	*1·9*	*2·0*	*2·4*	*2·7*	*3·9*	*6·0*	*12·3*	*54·7*	*100·0*
2,500	3,000		8,066	1,358	1,388	1,549	1,664	2,310	3,106	5,479	8,560	27,461	60,941
		Per cent.	*13 2*	*2·2*	*2·3*	*2·5*	*2·7*	*3·8*	*5 1*	*9·0*	*14·1*	*45·1*	*100·0*
3,000	4,000		9,559	1,755	1,831	2,030	2,443	3,246	4,496	6,050	9,506	24,672	65,588
		Per cent.	*14·6*	*2·7*	*2·8*	*3·1*	*3·7*	*4·9*	*6·9*	*9 2*	*14 5*	*37·6*	*100·0*
4,000	5,000		5,231	1,055	1,102	1,202	1,677	1,893	2,356	3,031	4,189	10,473	32,209
		Per cent.	*16·3*	*3·3*	*3 4*	*3·7*	*5·2*	*5·9*	*7 3*	*9·4*	*13·0*	*32·5*	*100·0*
5,000	6,000		3,278	660	675	880	1,105	1,221	1,296	1,654	2,181	5,405	18,355
		Per cent.	*17·9*	*3 6*	*3 7*	*4·8*	*6·0*	*6·7*	*7·1*	*9·0*	*11·8*	*29·4*	*100·0*
6,000	8,000		3,558	778	812	996	1,087	1,186	1,310	1,563	2,106	4,722	18,118
		Per cent.	*19·6*	*4·3*	*4·5*	*5·5*	*6 0*	*6·6*	*7·2*	*8 6*	*11·6*	*26·1*	*100·0*
8,000	10,000		1,822	372	491	581	556	576	595	639	914	1,970	8,516
		Per cent.	*21·4*	*4·4*	*5·8*	*6 8*	*6·5*	*6·8*	*7·0*	*7·5*	*10·7*	*23·1*	*100·0*
10,000	12,000		1,084	251	300	324	292	298	269	339	429	1,013	4,599
		Per cent.	*23·6*	*5·5*	*6·5*	*7·1*	*6·3*	*6·5*	*5·8*	*7·4*	*9·3*	*22·0*	*100·0*
12,000	15,000		903	245	248	240	220	221	216	272	300	771	3,636
		Per cent.	*24 8*	*6·7*	*6·8*	*6 6*	*6·1*	*6·1*	*5·9*	*7·5*	*8·3*	*21·2*	*100·0*
15,000	20,000		782	202	201	191	145	166	150	150	214	504	2,705
		Per cent.	*28·9*	*7·5*	*7·4*	*7·1*	*5·4*	*6·2*	*5·5*	*5·5*	*7·9*	*18·6*	*100·0*
20,000			928	197	168	144	114	104	105	120	170	403	2,453
		Per cent.	*37·9*	*8·0*	*6·8*	*5·9*	*4·7*	*4·2*	*4·3*	*4·9*	*6·9*	*16·4*	*100·0*
All ranges			45,817	8,534	8,856	9,854	11,421	13,608	17,308	24,507	39,238	125,047	304,190
		Per cent.	*15·1*	*2·8*	*2·9*	*3·2*	*3·8*	*4·5*	*5·7*	*8·1*	*12·9*	*41·0*	*100·0*

Source: 100th Report of the Commissioners of Inland Revenue. H.M.S.O. Cmnd. 341.

the rates of surtax change. This first table then is given by the vertical classification by incomes together with the figures in the last column (excluding those in italics). In effect, the first line of such a table would read: £2,000-2,500: 87,070 incomes.

The second suggested table could be made up from the first horizontal line, reading 0-10 per cent., and so on, together with the numbers (ex. italics) along the base. The first line of such a table would then read: 0-10 per cent.: 45,817 incomes; meaning that in each of those incomes up to 10 per cent. of any individual income was 'earned'. In the terminology of the Inland Revenue, this term applies to income or profits derived from an employment, trade, or profession. The balance of the income is 'unearned', meaning that it represents dividends and interest received on invested capital.

The third possible table can be formed by using the detailed analysis in the body of the table. For example, a distribution based on the proportion 'earned' in any particular income class, *i.e.*, the horizontal line relating to, say, £3,000-£4,000; indicating that 4,496 such incomes are earned to the extent of between 60-70 per cent., and so on. A similar analysis using any of the vertical columns is equally possible. Finally, in addition to all these suggested tables in which the absolute number of incomes are given, tables utilising the appropriate percentage figures may be constructed. The value of the percentages, as the table stands at present, lies in the fact that it becomes possible to compare the distribution of the earned proportions within the various classes of income (horizontally).

For example, in the column 30-40 per cent there are 324 incomes of between £10,000 and £12,000 and only 191 incomes of between £15,000 and £20,000. Yet the first figure represents the same proportion of all incomes in that particular class by size as does the second. They are both 7·1 per cent. The student reader may care to prepare such abbreviated tables as have been suggested and drawing such conclusions as may appear relevant. One criticism that can be made of the table, however, is that the percentage class limits are indeterminate. It will be noted that if the ratio of earned income to total income fell exactly on the limit, *e.g.* 10%, that particular income could be counted in either the 0%–10% class or the 10%–20% class. It may be argued that there will be relatively few incomes coinciding exactly with these

TABLE 14

DISPOSALS OF INDIGENOUS AND IMPORTED COAL FOR INLAND CONSUMPTION BY GRADE AND MAIN CONSUMER IN 1957

	Thousand tons										
	Large		Graded		Slacks and Smalls		Other Coal		Anthracite		TOTAL
	Tons	%	Tons	%	Tons	%	Tons	%	Tons	%	Tons
Electricity Works	8	*0·0*	818	*1·7*	44,329	*94·5*	1,223	*2·7*	513	*1·1*	46,890
Gas Works	2,672	*10·1*	16,876	*63·8*	474	*1·8*	6·440	*24·3*	3	*0·0*	26,465
Coke Ovens	691	*2·2*	1,676	*5·2*	27,076	*84·3*	2,667	*8·3*	—	—	32,110
Railways	11,261	*97·9*	126	*1·1*	83	*0·7*	28	*0·2*	7	*0·1*	11,505
Industry .. (*a*)	3,562	*8·8*	16,116	*39·7*	19,693	*48·5*	532	*1·3*	693	*1·7*	40,596
Domestic Merchants ..	27,652	*88·6*	2,704	*8·7*	67	*0·2*	98	*0·3*	691	*2·2*	31,212
Miscellaneous .. (*b*)	3,191	*27·0*	3,703	*31·3*	3,695	*31·2*	680	*5·7*	565	*4·8*	11,834
TOTAL	49,037	*24·4*	42,019	*21·0*	95,416	*47·6*	11,668	*5·8*	2,472	*1·2*	200,612

Source: Ministry of Power Statistical Digest (*H.M.S.O. 1958*).

(*a*) including establishments of annual consumption of less than 1,000 tons;

(*b*) excluding N. Ireland Service Depts., Waterworks and non-industrial establishments and coastwise bunkers.

limits but it still does not alter the fact that this constitutes an error in defining the class limits.

Another example of a neat tabulation is given in Table 14, taken from the Ministry of Power's Statistical Digest. It is designed to classify the various grades of coal sold according to the main home consumer groups. It is subdivided to illustrate the proportions of the total consumption of each main consumer group of the various grades of coal (the percentages add up to 100 crosswise). It shows in addition the actual tonnages consumed which brings out such facts that industry, although consuming only about 20 per cent of total small coal used, purchases this grade for nearly half its total requirements.

A more unusual form of table is shown in Table 15 taken from the *Ministry of Labour Gazette* for August 1958. It shows an analysis of answers to questionnaires sent to nearly 7,800 people who entered Industrial Rehabilitation Centres in 1956 and completed their courses there. Six months after they had completed their courses they were asked by means of this questionnaire whether they were satisfactorily placed in employment. Those who were recorded as being 'in training' were in fact taking a further course under the Ministry's Vocational Training scheme. The table itself provides an example of the way in which almost any type of numerical data can be portrayed in a clear-cut fashion.

One interesting lesson that can be learnt from Table 15 is the need to enquire closely into the definition of categories in tables, particularly if they are of a qualitative nature, even though they may appear quite clear at first sight. Thus it appears from the table that the medical groups that had the greatest difficulty in retaining employment (recorded in the column 'Not in employment at date of enquiry but some work since course') were those suffering from mental or nervous disorders; with the notable exception of those in the 'able-bodied' group. However, on closer investigation, one finds that although these people were sent to the Rehabilitation Units as being 'able-bodied' upon medical examination after arrival nearly 90 per cent were found to have some disability which in half the cases was of a mental or nervous nature. Therefore the high proportion of able-bodied who had difficulty in retaining employment is not inconsistent with the general concept that it is the mental or nervous type of

TABLE 15

FOLLOW-UP OF PEOPLE WHO ENTERED INDUSTRIAL REHABILITATION UNITS IN 1956 AND COMPLETED THEIR COURSE

Medical Group	Number of replies to follow-up enquiry	Analysis of replies expressed as percentages of total replies in each medical group					
		In training	In employment and satisfied	In employment but not satisfied	Not in employment at date of enquiry but some work since course	No employment by date of reply	
						Unemployed	Sick
Amputations	154	19·5	44·2	3·2	14·3	14·2	4·6
Arthritis and Rheumatism ..	265	12·8	42 6	5·7	12·1	14·4	12·4
Dieseases of:							
Digestive system	233	13·3	50·2	8·2	12·4	13·2	2·7
Heart or Circulatory system	486	11·7	45·7	6·8	15·0	15·4	5·4
Respiratory system (not T.B.)	402	8·7	48·3	7·0	14·9	11·2	9·9
Skin	71	16·9	45·1	8·4	11·3	15·5	2·8
Ear defects	57	12·3	40·4	10·5	19·3	17·5	—
Eye defects	86	16·3	48·8	8·1	10·5	13·9	2·4
Injuries and diseases of:							
Lower limbs	337	13·0	54·9	8·6	12·5	6·8	4·2
Upper limbs	267	13·1	56·2	7·5	10·9	9·8	2·5
Paraplegia	51	15·7	37·2	—	15·7	21·6	9·8
Other spinal disorders ..	291	16·8	47·8	5·2	12·0	13·4	4·8
Epilepsy	231	13·0	39·8	6·1	21·6	11·3	8·2
Other organic nervous diseases	357	10·1	41·7	4·5	14·3	18·8	10·6
Psychoneurosis	694	10·9	43·8	9·1	18·2	13·0	5·0
Mental deficiency	102	2·0	46·0	4·9	18·6	22·5	5·9
Psychosis	236	12·7	43·2	4·7	22·5	10·2	6·7
Respiratory Tuberculosis ..	1,656	28·3	45·3	5·8	10·1	8·1	2·4
Other Tuberculosis	171	23·4	44·4	5·3	9·9	13·5	3·5
All other disabilities	584	15·9	42·0	6·5	14·8	13·6	7·2
Able-bodied	41	12·2	43·9	2·4	24·4	17·1	—
All groups	6,772	16·8	45·6	6·5	13·8	12·2	5·1

Source: Ministry of Labour Gazette, August 1958.

person who had the greatest difficulty in retaining employment. It may be of interest to consider the effects of other factors on people using these units, *e.g.* comparative youth of the cases of respiratory T.B., previous education and employment, unsettled compensation cases, etc.

Table 16 is a slightly more complex example of a classification and analysis used in the preceding table. The reader should note the initial classification of the women according to whether they lived in textile areas or elsewhere, since it may be assumed that their attitudes will be different from one another. The next most important sub-classification is whether the woman was single or married; clearly an important consideration when trying to recruit women into industry. Finally, these groups are again classified as 'experienced' or 'inexperienced'. It will be apparent that this 3-stage classification might well have been differently arranged. For example, the table might have started by classifying the women as married or single, then dividing these two groups according to whether they lived in a northern area or not and then by their experience. The actual classification employed depends on what characteristics are the most significant; in this case it is clear that the woman's background of life in a textile area or elsewhere was a more important factor than anything else. Note too the frequency of the reason 'Health grounds' given for not entering the industry; this almost certainly was frequently used to conceal the real attitude to work in the textile industry.

The final example, Table 17, is included as an illustration of an 'over-loaded' table. The data in this case are far more difficult to extract than in the previous examples, due to the mass of detail which has been introduced. By sectional rulings and sub-totals, the reader is enabled to trace a path through the mass of data, but the table provides a useful indication of the probable limits of human capacity to grasp large quantities of data. It would have been helpful if the main heads had been extracted and formed into a smaller summary table, thereby enabling the main features to be picked out more easily. It would also have helped if the money totals had been rounded to the nearest ten thousand.

The foregoing examples of tables covering a wide variety of data from the economic, industrial and social fields illustrate how well a great deal of information may be simply presented. There is hardly a single case where data are not more

TABLE 16

EAST AND WEST RIDING REGION

Table showing the Reasons advanced by Unemployed Women for not desiring to return to, or take up Employment in the Woollen and Worsted Textile Industry. June 1946.

		Textile Areas						Non-Textile Areas					
		Married		Single		Total		Married		Single		Total	
		Experi-enced	Inexperi-enced	Experi-enced	Inexperi-enced	Experi-enced	Inexperi-enced	Experi-enced	Inexperi-enced	Experi-enced	Inexperi-enced	Experi-enced	Inexperi-enced
Number of apparently suitable women interviewed	1	247	73	154	38	401	111	49	520	16	710	65	1,230
Number included in (1) who accepted employment in Wool Textiles	2	75	14	57	11	132	25	10	82	7	75	17	157
Number included in (1) who refused to consider employment in Wool Textiles	3	172	59	97	27	269	86	39	438	9	635	48	1,073
Analysis of Reasons advanced for refusing Employment as at (3):	4												
Rates of wages unsatisfactory	a	19	10	4	4	23	14	4	13	2	9	6	22
Working conditions in mills unsatisfactory	b	21	10	20	11	41	21	4	6	1	—	5	6
Shorter hours of work required	c	24	7	1	—	25	7	4	16	1	5	5	21
Part-time work only required	d	29	10	3	—	32	10	—	5	—	—	—	5
Shift work (6 a.m. to 2 p.m.; 2 p.m. to 10 p.m. required)	e	1	—	—	—	1	—	—	—	—	—	—	—
Five-day working week required	f	2	—	—	—	2	—	—	—	—	—	—	—
Objection to early morning start	g	9	3	—	—	9	3	4	8	1	8	5	16
Objection to working away from home	h	—	—	—	—	—	—	12	301	4	531	16	832
Health grounds	i	50	5	38	2	88	7	—	4	—	9	—	13
Work too heavy	k	1	—	—	—	1	—	—	—	—	—	—	—
Work too noisy	l	1	—	—	—	1	—	—	—	—	—	—	—
Does not like mill work	m	5	10	28	1	33	11	4	1	—	67	4	68
Wants to join W.L.A.	n	—	—	—	1	—	1	—	—	—	—	—	—
Wants outdoor work	o	2	—	—	—	2	—	—	—	—	—	—	—
Wants office work	p	7	2	1	2	8	4	—	—	—	—	—	—
Afraid of machinery	r	—	1	—	4	—	5	—	—	—	—	—	—
Prefers engineering work	s	1	—	—	1	1	1	—	—	—	—	—	—
Wants nursing	t	—	—	1	—	1	—	—	—	—	—	—	—
Not enough prospects	u	—	—	1	1	1	1	—	—	—	6	—	6
Poor canteen facilities	v	—	—	—	—	—	—	5	—	—	—	5	—
No morning or afternoon break	w	—	1	—	—	—	1	2	—	—	—	2	—
Too old	x	—	—	—	—	—	—	—	84	—	—	—	84

Source: Ministry of Labour

TABLE 17 VALUE OF OUTPUT OF ELECTRICITY

	£ thousand					
	Generation and Main Transmission		Distribution		Total	
	1955	1956	1955	1956	1955	1956
Sales of electricity:						
To consumers (*a*)	6,884	8,931	381,829	438,566	388,713	447,497
Within the industry ..	262,020	302,279	262,020	302,279	—	—
Steam and hot water ..	691	781	—	—	691	781
Ashes	185	269			185	269
Scrap metal	202	267	1,376	1,508	1,578	1,775
Other out sold	13	9			13	9
Total	269,995	312,536	121,185	137,795	391,180	450,331
Rents: Meter rents ..			379	365	379	365
Hire of appliances ..			2,654	2,696	2,654	2,696
Total			3,033	3,061	3,033	3,061
Proceeds from sale of purchased appliances and sale of reconditioned appliances withdrawn from hire (*b*)			35,046	28,117	35,046	28,117
Less Cost of purchased appliances sold and value (before reconditioning) of appliances withdrawn from hire			27,193	20,997	27,193	20,997
Payment for renovation of appliances by other firms			75	88	75	88
Net proceeds from resale of appliances			7,778	7,032	7,778	7,032
Work charged for:						
Installation and maintenance of public lamps			2,730	3,052	2,730	3,052
Fitting and maintenance of wiring and appliances ..			11,966	13,520	11,966	13,520
Other work charged to consumers	91	83	4,811	5,519	4,902	5,602
Total	91	83	19,507	22,091	19,598	22,174
Other work done by industry's employees (value of materials and wages):						
On depots, workshops, offices and other buildings:						
New construction (including extensions)	937	1,070	1,379	1,280	2,316	2,350
Repairs and maintenance ..	*1,102*	*1,379*	*2,861*	*3,239*	*3,963*	*4,618*
On plant and machinery:						
New construction (including complete renewals) ..	2,249	2,402	13,523	11,321	15,772	13,723
Repairs and maintenance .	*12,883*	*14,334*	*9,810*	*10,524*	*22,693*	*24,858*
On mains and services:						
New construction (including extensions and complete renewals)	415	538	38,180	37,165	38,595	37,703
Repairs and maintenance .	*539*	*541*	*6,560*	*7,336*	*7,099*	*7,877*
Total new constructional work done (*c*)	3,772	4,067	53,748	50,402	57,520	54,469
Total repair and maintenance work done (*c*) .. .	*14,694*	*16,399*	*19,421*	*21,342*	*34,115*	*37,741*
Total value of output and work done (excluding *repair and maintenance work*)	273,858	316,686	205,251	220,381	479,109	537,067
Less Cost of purchased material and fuel used .	157,702	187,038	44,059	44,712	201,761*d*	231,750
Less Transport costs for carriage of goods outwards	52	46	140	133	192	179
Net output	116,104	129,602	161,052	175,536	277,156	305,138

Source: Ministry of Power Statistical Digest 1957, Table 74 (H.M.S.O. 8958)

(*a*) The amount chargeable for the electricity actually supplied excluding rents shown separately.
(*b*) Includes sales on extended credit terms of £23,670 thousand and £16,101 thousand in 1955 and 1956 respectively.
(*c*) Includes value of work done between Divisions and Area Boards, details of class of work are not available.
(*d*) Includes £19,264 thousand in 1955 and £21,700 thousand in 1956 for carriage of goods inwards

satisfactorily presented by tables than by the method of linking the figures by a running commentary. No one who has listened to a speech or read a report crammed full of so-called statistics will deny that many headaches could have been saved by a simple set of figures with a separate text.

Summary

The construction of statistical tables is simple enough, but it may help the reader if the main points concerning tabulation are recapitulated:

1. The title of the table should be concise and self-explanatory.
2. The table should not be overloaded with detail.
3. If the table is complex and cannot be broken down suitably, thick and thin rulings, heavy printed sub-titles and coloured inks will all help in clarifying the overall picture.
4. Columns of figures which are directly comparable should be kept together, as should percentages or ratios relating to the absolute quantities.
5. If columns are to be aggregated and the totals are significant for comparison, they can be put at the top of the columns rather than below.
6. Units of measurement should be clearly shown, and if necessary defined. The same applies to column headings such as those to which attention was drawn in the text, *e.g.*, Table 13.

How far the examples given in the preceding pages meet these requirements, the reader may judge for himself. A little experimenting to improve the layout of, say, Table 17 may help to drive home the main lessons! If any reader wishes to pursue this topic further, the annual reports of the nationalised industries and public corporations provide a prolific source of information in tabular form. A study of the tables illustrating the text of Chapters XIV and XVII should also prove helpful in this respect.

In conclusion, it may be emphasised that tabulation is not a dull task to be allocated to someone who can merely count. The final tables should be drafted only after every consideration has been given to the nature of the data, the purpose of the enquiry, and the way in which this stage of the enquiry may best be

covered to facilitate further work. Nothing, and least of all a statistical enquiry, is perfect, but there is no reason for creating further difficulties by indifference to simple but fundamental points in tabulation.

CHAPTER V

GRAPHS AND DIAGRAMS

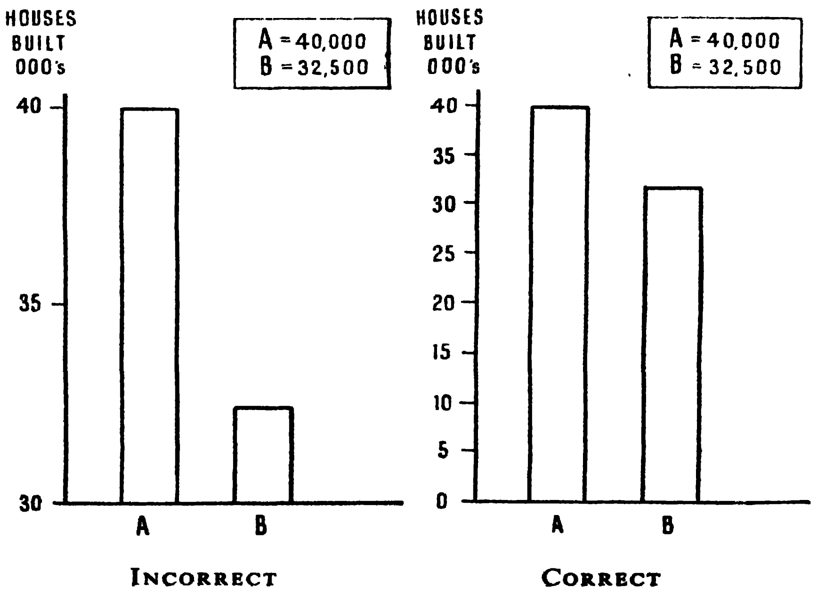
However informative and well designed a statistical table may be, as a medium for conveying to the reader an immediate and clear impression of its content, it is inferior to a good chart or graph. Many people are incapable of comprehending large masses of information presented in tabular form; the figures merely confuse them. Furthermore, many such people are unwilling to make the effort to grasp the meaning of such data. Graphs and charts come into their own as a means of conveying information in easily comprehensible form. It is for such reasons that the government has produced popular versions of important White Papers in the form of multi-coloured booklets full of charts and simple figures. Such pictorial representation admittedly reduces the amount of detail that can be put across to the reader, but very often it is not the detail which is important, but rather the overall picture. For example, few citizens can give figures of the extent of this country's post-war balance of payments position, but most of them have been made aware by publicity employing charts that an expansion of exports is still necessary to pay for our foodstuffs and raw materials.

Diagrammatic representation of statistical facts is not only popular with the lay public; it is also extremely useful to the statistician. For example, a few well designed but simple charts showing the trend of sales and costs will be infinitely more eloquent at a board meeting than a mass of detailed monthly figures. Even the statistician himself will employ diagrams to ascertain the pattern or distribution of his data because the character of the distribution will sometimes determine the type of statistical analysis he will employ. There is a large number of diagrammatic forms to choose from; some of the most popular types of chart are reproduced in this chapter. The variety does not arise because statisticians as a class are particularly artistic; the data will usually determine the type of chart used. While there are certain obvious rules regarding the construction of charts, the most important consideration is commonsense. A

good policy to adopt is to consider the finished diagram and ask what conclusions can be drawn from it. If they differ substantially from the impressions derived from a brief study of the actual data, then the chart should be scrapped. Some loss of detail is inevitable, but the chart need not be misleading.

A good illustration of poor design is given in Figures 1 and 2 below. Some years ago during a municipal election, one party anxious to impress upon the electorate its superior performance in house building put up a poster on the hoardings on the lines of the left hand part of Figure 1. By not drawing the base line upon which the vertical bars were drawn from zero, the relative performance of that party was greatly enhanced in the eyes of the casual observer. The fact that the correct figures were inserted in the chart probably did little to counter the first impression. The correct method of drawing this chart is given on the right hand side of the Figure 1. Some criticism can also be directed against the left hand chart in Figure 2, which illustrates an advertisement used by one newspaper to demonstrate its

Figure 1



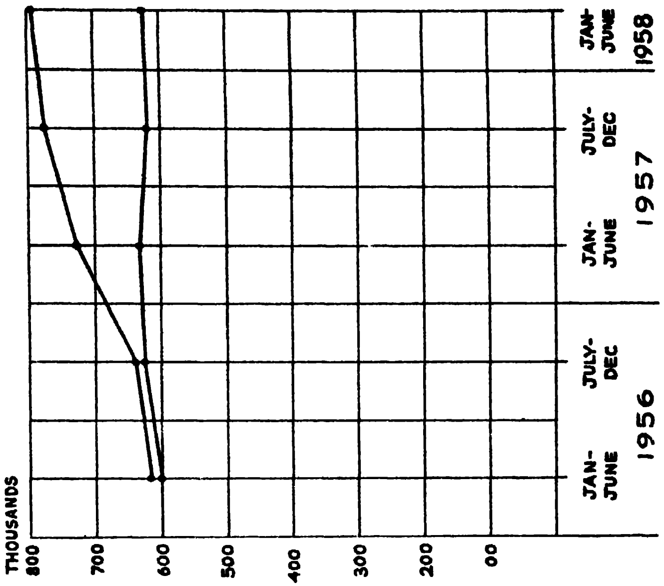
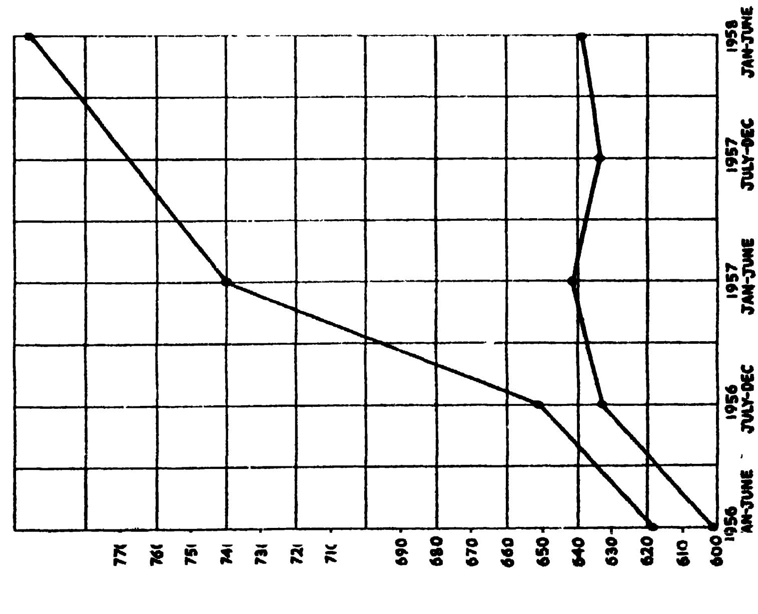

THOUSANDS
800
700
600
500
400
300
200
00
JAN-JUNE
JULY-DEC
1956
JAN-JUNE
JULY-DEC
1957
JAN-JUNE
1958

Figure 2


690
680
670
660
650
640
630
620
610
600
1956 JAN-JUNE
1956 JULY-DEC
1957 JAN-JUNE
1957 JULY-DEC
1958 JAN-JUNE

popularity *vis-à-vis* its main rival. The vertical axis is clearly marked with the actual circulation figures, but once again, by omitting the base line and the entire lower part of the chart the performance of the advertiser's paper is greatly enhanced. It is undoubtedly true that the one paper has in the space of two years outstripped its rival, but by using a different scale and redesigning the chart, the picture can be made to look rather different. The reader should compare the right hand chart with the original on the left. This is also a bad chart, but for different reasons. It is badly designed with the bulk of the space wasted. As an exercise the student should draw a new graph. The scale on the left-hand graph is sufficiently detailed to enable approximate figures of circulation to be extracted. In this new graph, the vertical axis should show the origin and at a point slightly above it show a distinctive break as in Figure 9 (page 75) up to the first figure of 600,000 from which point the scale can then be marked off.

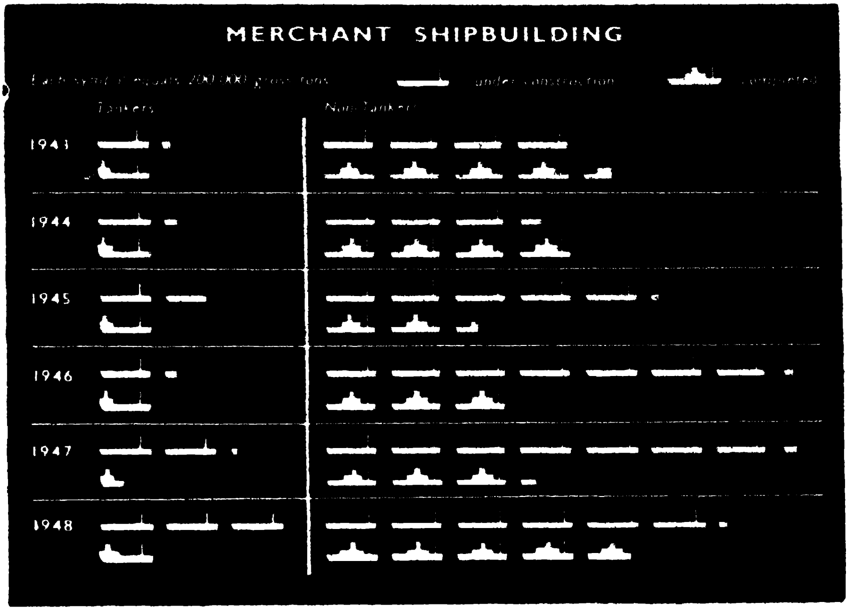
Pictograms

Before discussing the various types of diagrams and their uses, a clear distinction should be drawn between the highly simplified and sometimes coloured pictorial diagrams, such as are employed by the Government Departments to explain the economic situation as well as by some leading companies to bring out the main features of their development in the past year, to supplement the Chairman's speech and the graphs employed in statistical work proper. Within limits, the former type, 'pictograms', as they are sometimes called, are most useful. Fig. 3*a* illustrates the annual tonnage of ships and tankers built and under construction in each of the years from 1943-48. The relationship between the years for the two sets of figures, *i.e.*, tankers and non-tankers, is indicated by the *number* of little ships, each of which represents 200,000 gross tons.

Comparisons of this kind are often made by drawing the larger quantity as a larger version of the smaller, *e.g.*, one large ship for 1948, and a much smaller one for 1943. The difficulty then arises as to how the relationship is to be determined. Is it by area, *i.e.*, two dimensions, or is it by volume, *i.e.*, three dimensions? In any case, the visual effect is generally inferior to the former method. This method is employed in Fig. 3*b*, which illustrates

Figure 3a

PICTOGRAMS



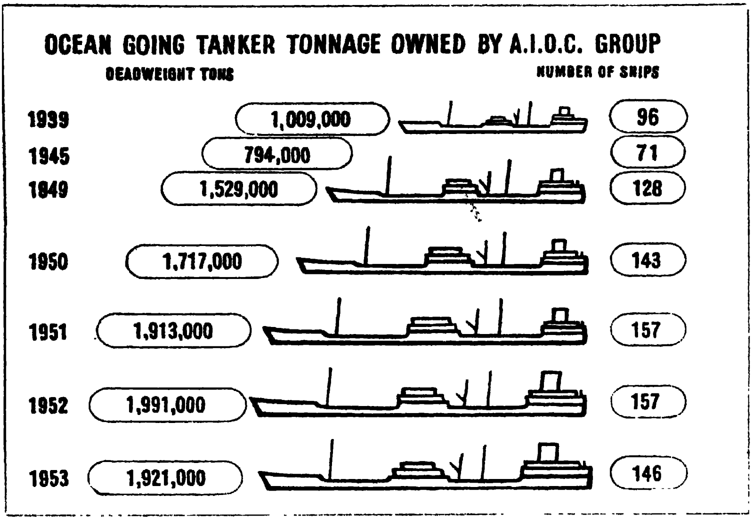
This diagram is reproduced by kind permission of the editor of 'Barclays Bank Review'.

the post-war growth of the tanker fleet of the Anglo-Iranian Oil Company (now known as the British Petroleum Co.) in the period 1939-53. The weakness cited above concerning the comparison between the relative sizes of the tankers is overcome by inserting the actual figure against each drawing. In fairness to those responsible for the original drawings, it may be mentioned that their effectiveness was greatly enhanced by the use of colour. Even in black and white, however, the story they tell is simple and clear.

Pie Charts

One of the simplest methods of illustrating the distribution of a particular population or aggregate is the *pie* diagram. This consists of two circles, preferably of equal area, divided into sectors. Figure 4 shows the age distribution of the population of England

Figure 3b



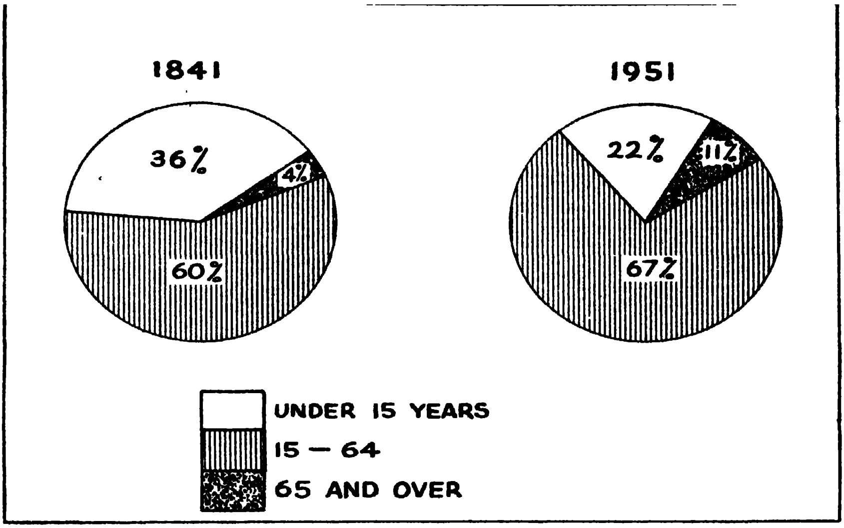
This diagram is reproduced by kind permission of the British Petroleum Co Ltd.

and Wales in 1841 and 1951. Within each circle there are only three sectors, the meaning of which is indicated in the key to the diagram. Sometimes such diagrams are drawn with a number of quite small sectors but this is not good practice. The essence of such diagrams is simplicity, three or four sectors are quite enough for the eye to comprehend. Opinions vary as to whether the percentage figures should be inserted within the diagram to show the relative size of the constituent sectors. Where there are at most four sectors the figures do clarify the picture, but if there are more than this number and sometimes the percentage has to be indicated by an arrow outside the circle to the circumference of the sector, the figures only tend to confuse.

Occasionally, the two circles represent different aggregates and are therefore drawn with different areas. For example, the 1951 circle could have been drawn with an area about four times as large as that for 1841 to indicate that the population has expanded fourfold. This again is not really good practice since the reader has then to consider two aspects, the relative size of the

Figure 4

AGE DISTRIBUTION OF POPULATION IN ENGLAND AND WALES AT CENSUS DATES 1841 AND 1951.



Sources: Registrar General's Review.

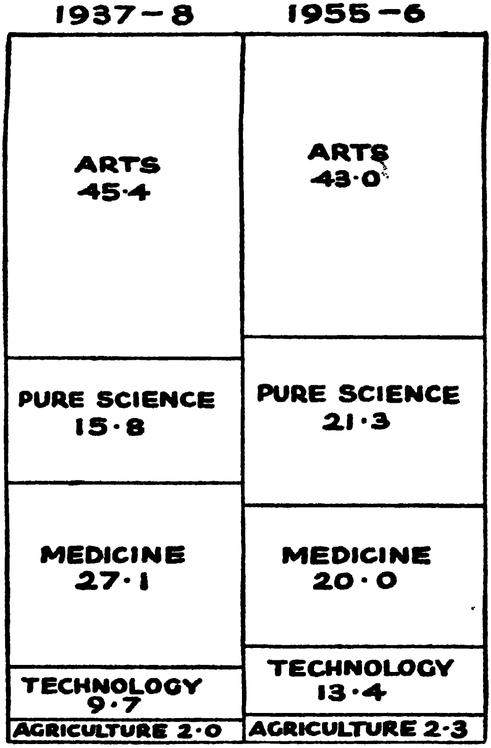
circles which is quite difficult to comprehend, and the distribution of their areas among the sectors. It is a good rule to keep this type of diagram as simple as possible; if details are necessary then another diagram bringing out different aspects of the data should be drawn. For example, we could illustrate the fourfold expansion of the population during the period by drawing two vertical bars or columns as in Figure 1, that for 1841 being drawn about one-quarter the height of that for 1951. Then the pie chart as shown in Figure 3 can be used solely to illustrate the change in the age structure of the population. The actual construction of the pie diagram is quite easy. Each percentage is converted into its equivalent part of 360 degrees; thus 60 per cent of 1841 requires a sector of 216 degrees, *i.e.*, 60/100 × 360.

Bar Charts

A simple variant of the pie chart is provided by what is sometimes referred to as a block diagram or bar chart. Figure 5 shows the distribution of students in full-time attendence at universities

Figure 5

PERCENTAGE DISTRIBUTION OF FULL-TIME UNIVERSITY STUDENTS IN GREAT BRITAIN BY FACULTIES.

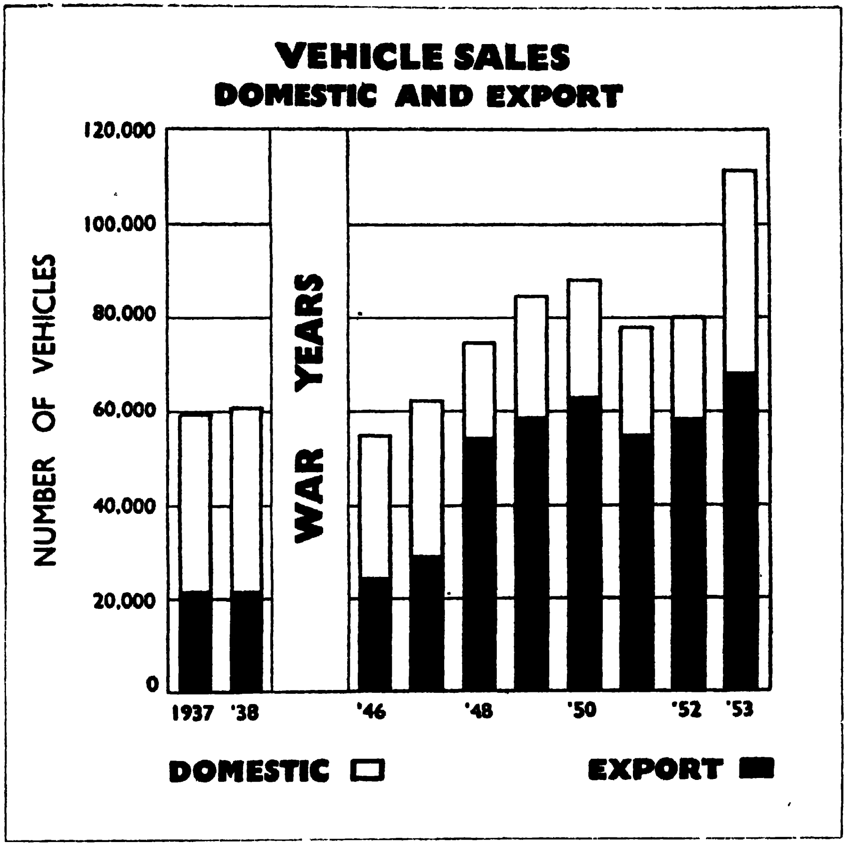


Source: University Grants Committee Report.

in Great Britain according to the course they are following. The main point of the diagram is to bring out the changed distribution and for this reason the two blocks or bars, the actual width is not really important, are contiguous. This facilitates the interpretation of the figures. As on occasion with the pie diagrams, the two blocks are sometimes drawn to bring out the absolute change in numbers. For example, there were 49,000 full-time students in 1938-9 but 85,000 in the academic session 1955-6. This point could have been brought out by drawing the second block approximately twice as high as the first, but the

Figure 6

COLUMN DIAGRAM



Reproduced by kind permission from the 1953 Report of the Chairman of Vauxhall Motors Limited.

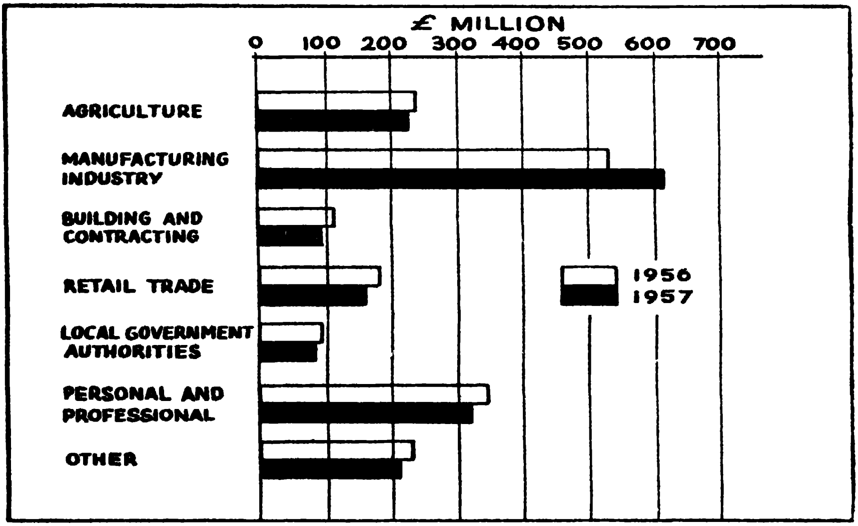
effectiveness of the comparison of the distribution of students between faculties would have suffered. Like most human beings, charts function better if they try and do one thing at a time!

Figure 6 illustrates one of the most frequent uses of the bar or column diagram. It depicts a time series, in this case the annual sales of Vauxhall Motors Ltd. during the period 1937 to 1953, omitting the war years. The division of each bar to indicate the shares of annual output sold at home and abroad is quite simple.

The chart is good since it tells its story simply; the eye can assimilate all the information quite easily. An alternative form of bar diagram is given in Figure 7 which illustrates the distribution of bank advances and their amount for certain sectors of the economy. At the same time the bars for 1956 and 1957 are contiguous so that the absolute change in the volume of bank lending to each sector can be easily noted. The horizontal scale across the top of the diagram enables a rough assessment to be made of the volume of advances outstanding with the industries

Figure 7

BANK ADVANCES - SELECTED CATEGORIES IN 1956 AND 1957.

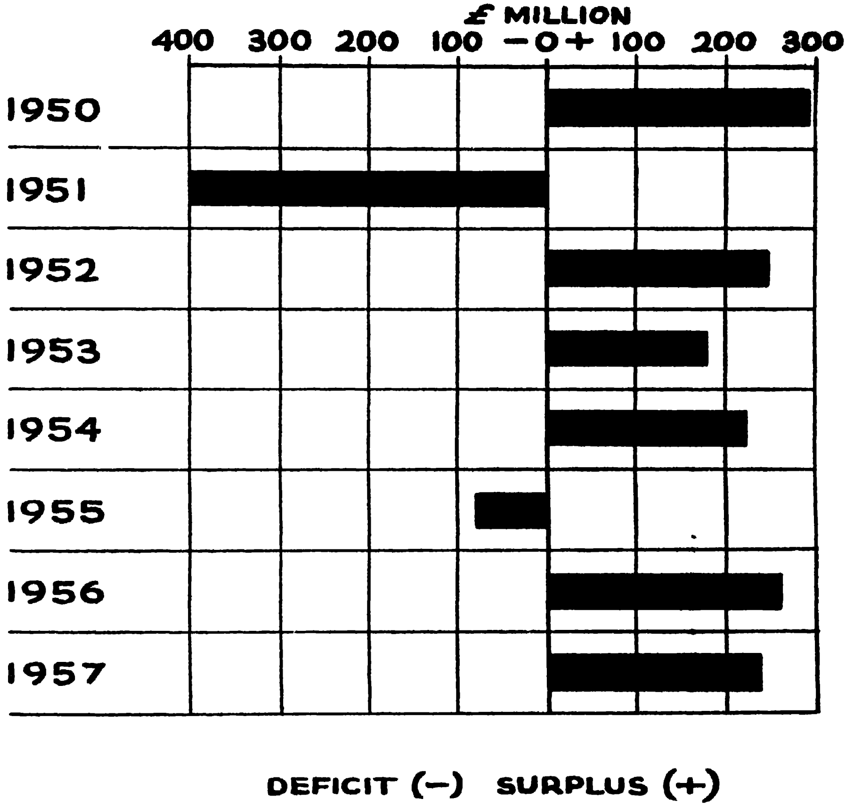


Source: Economic Trends.

shown. When the quantities to be illustrated are both positive and negative, *e.g.*, annual profits and losses, the bar diagram can be adapted in the manner shown in Figure 8. For each year between 1950 and 1957 the surplus or deficit on the United Kingdom's overseas current account can be read off against the horizontal scale at the top of the diagram. The visual impression is also good; the distinction between surplus and deficit years is immediately apparent. In this case the bars are drawn horizontally; the same effect can be achieved by drawing vertical bars above and below a horizontal line marking the origin.

Figure 8

U.K. BALANCE OF PAYMENTS CURRENT ACCOUNT ANNUAL SURPLUSES AND DEFICITS 1950-7.



Source: Economic Survey 1958.

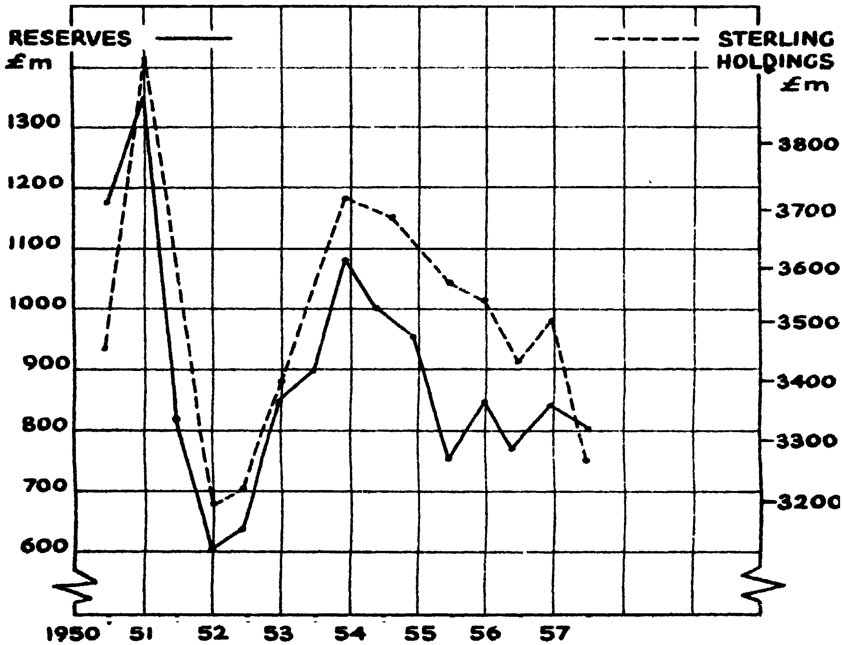
Charts for Time Series

The graph for depicting time series is extensively employed, especially in business or any other sphere where data are collected over periods of time. Most readers will recall the principles underlying the construction of a graph as taught them at school, but some may be glad to have their memories refreshed. The vertical axis on the left is called the ordinate, and the horizontal

base line the abscissa. It is customary to plot the time factor along the abscissa; for it is the *independent* variable. The variable factor is plotted along the ordinate; it is known as the *dependent* variable because it varies from year to year as compared with the year as a unit of time which is quite unalterable and 'independent' of other factors. The position on the graph of any point is located by the co-ordinates of that point. Thus, if we use Figure 9 for illustration, it is clear that in June 1952 the gold and dollar reserves were some £600 million. This point is located on the graph by the intersection of a vertical line drawn from the abscissa where it is marked mid-1952, and a horizontal line from the point on the ordinate marked £600 million.

Figure 9

U.K. GOLD AND DOLLAR RESERVES AND OVERSEAS COUNTRIES' STERLING HOLDINGS 1950-7.



Source: Economic Survey 1958.

Both the data depicted in Figure 6 as well as those in Figure 9 are time series. In the first diagram the data for each year were depicted by a bar (it could equally well have been a line) and in the second graph each half-year's figure is linked with the next to form a continuous line across the graph. In this case the line graph brings out very clearly the fluctuations in the monthly totals and in particular makes an effective means of comparing the movements in the gold and dollar reserves with those in the overseas sterling holdings. But, as a rule, when a time series is formed by quantities which are stated at a given point of the year, it is better to emphasise this point in the graph so that the graph is not used to read off mid-year totals which are assumed to be reasonably accurate. For example, we could plot the annual number of births for the decade 1945-54 by a line chart since births occur all the year round. But if annual figures for that decade were derived from a count on a particular day of the year, *e.g.*, bank advances as at June 30th, then, strictly speaking, the series should be depicted in the graph as a series of unconnected vertical bars at regular annual intervals along the base. The line graph for time series has many merits, not least for comparative purposes as in Figure 9. Therefore, the foregoing rule is often ignored, and at the base of the line chart is written, *e.g.*, totals at June 30th, or December 31st, or weekly average for year, to indicate the basis on which the totals have been measured.

The main interest in Figure 9 as an illustration of graphical methods lies in the use of two scales for the two series. The ordinate on the left is marked off in units of £100 million and the curve representing the reserves is read off against this scale. Because the sterling holdings of other countries greatly exceed the United Kingdom's reserves, if the curve representing fluctuations in their size had to be plotted against the same scale, then the left-hand ordinate would have to be greatly extended up to £3,800. The sterling holdings curve would then be stuck up at the top of the graph and there would be a great empty space between the curves. This could be reduced by breaking the left-hand ordinate so that it consisted of two separate scales, the lower one reading as at present £600 – 1,300 million and above it, broken as shown in Figure 2*a*, a scale reading £3,200 – 3,800 million. This would eliminate a large part of the 'empty space' created by using a continuous scale from £600 – 3,800 million

without any break. But it would still take up a lot of space which in the average sized graph would entail the reduction of the scale used. Instead of letting ½″ on the ordinate represent £100 million, it would have to represent (say) £500 million. Such a scale would then virtually obliterate the minor movements in the two values. In this connection note that the two scales both show a break above the origins so that they can be started at £600 and £3,100 million respectively. Unless such a break were shown it would be necessary, as was the case in Figure 2, to plot all the values from £0 upwards along the ordinates, again wasting space and reducing the scale of the graphs.

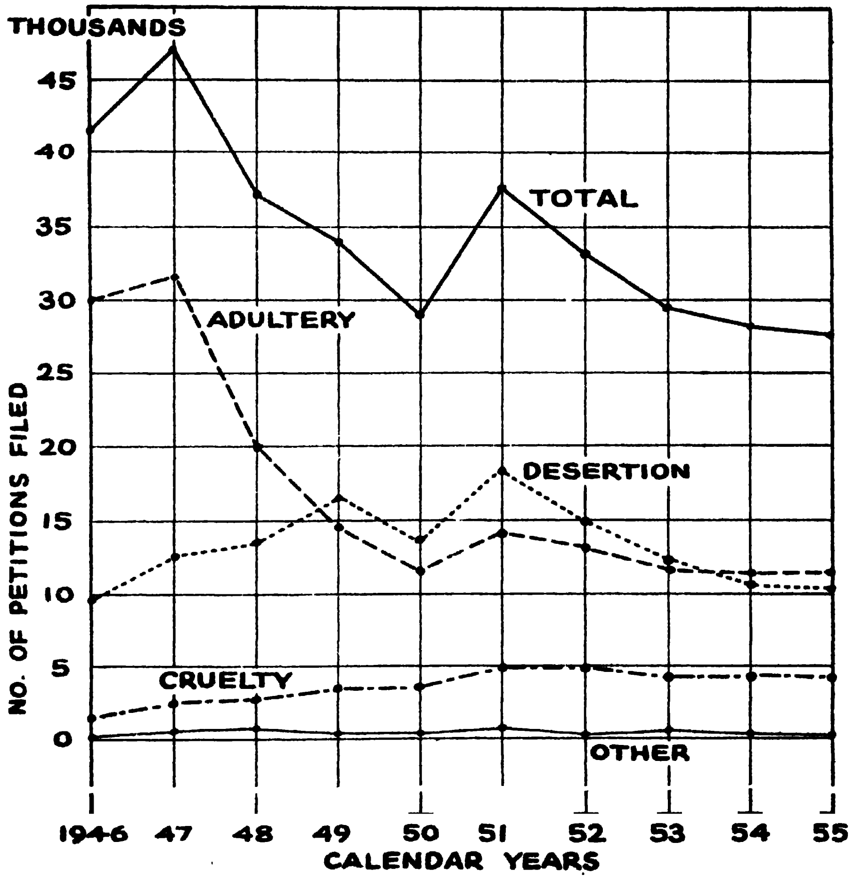
The method of plotting the two curves used in Figure 9 is extensively employed for all the above reasons. It entails the use of two ordinates on a common scale so that like movements in the two figures are shown by the same rise or fall on the graph. By using two ordinates the two curves can be almost superimposed upon one another thereby facilitating the comparison of their respective movements. Had the two curves intersected at more than one point, they would have been more difficult to interpret. In such cases, the right-hand scale can be lifted bodily up the graph an inch or two so that frequent intersection of the curves is avoided. The actual construction, however, of Figure 9 could be criticised on one point. The two scales are equal; *i.e.*, the same vertical distance on each equals a change of £100m. Such a change is, however, of very different *relative* significance to the two totals. For example, the drop of over £700m. in both figures in 1952 represents almost half the reserves, but barely twenty per cent. of the sterling holdings, although the conclusion drawn from the graph as drawn is that it is an equivalent fall – as it is, in *absolute* terms. On the other hand, if the object of the chart is to emphasise the inter-dependence of the reserves and overseas liabilities, it serves its purpose.

The data depicted in Figure 10 are the number of divorce petitions filed each calendar year. They are annual totals and can therefore be depicted by a line graph. In such cases the figure for each year is plotted against the middle of the interval described as a year along the abscissa. When a line graph is drawn for data given 'as on December 31st each year', then the point is usually plotted above the end of the interval occupied by the year along the base. In addition to the annual fluctuations in the total

number of petitions drawn in Figure 10, there are also lines representing the annual totals of petitions according to the grounds upon which the petition is based. It will be noted that adultery was by far the most important reason after the war but a few years later desertion became a more important cause. The sudden increase in the number of petitions in 1951 is due to a change

Figure 10

DIVORCE PETITIONS FILED ANNUALLY IN ENGLAND AND WALES 1946-55
(ANALYSIS BY GROUNDS OF PETITION)



Source: Annual Abstract 93. Table 86.

in the divorce law which facilitated divorce for certain persons, not to any sudden dislike of the married state! Cruelty is still relatively unpopular as a ground for divorce, although there seems to have been a small rise in the number of petitions on this ground. 'Other' causes are lunacy, presumed decease and rape, etc.; they are very infrequently pleaded. Note in this diagram how the abscissa has been arbitrarily 'dropped' below the origin (0) on the ordinate simply to enable the 'Other' curve to be put in and read more easily.

The main value of this type of graph is that it brings out any relationship which may or may not exist between the movements in the various values, as well as indicating the relative importance of one as against the other; in this graph the relative unimportance of cruelty as a ground for divorce and the extensively used ground of adultery. There is one danger with this type of graph; it is easy to overload it with several curves which intersect each other sometimes more than once, so that it becomes quite difficult to read and follow the trend of the curves. If it is necessary to compare several times series it is better to draw them with a common abscissa and break the ordinate up so that each line has its own scale. They can then be drawn one above the other. Provided the vertical scale is the same for all the values plotted, *e.g.*, ½″ to 500 petitions, £100 or whatever the dependent variable may be, any relationship between the series can be easily observed.

Logarithmic Scale Graphs

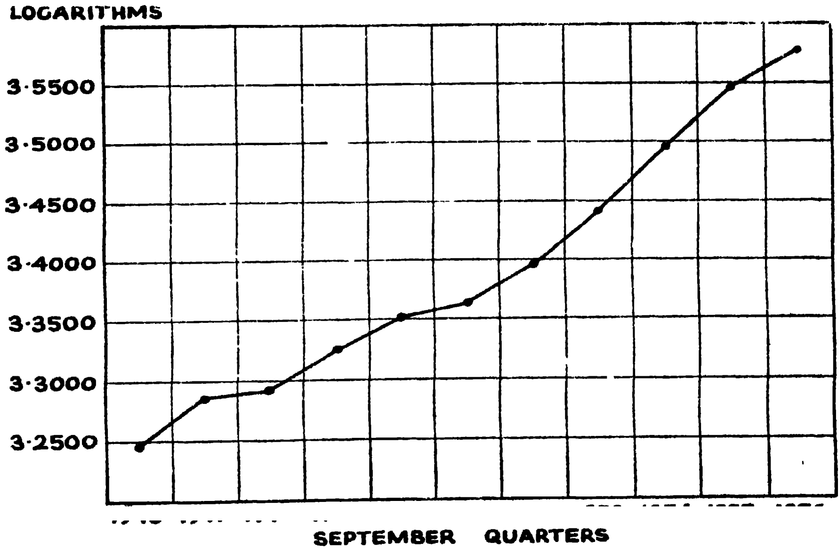
Changes in the annual totals depicted in Figures 9 and 10 were shown absolutely, in other words a movement of £100 million in the gold and dollar reserves took up the same space as a change of that amount in the sterling holdings. If the gold and dollar reserves rose from £1,000 to £1,100 the £100 increase would represent a rise of 10 per cent. If the sterling holdings rose from £3,300 to £3,400, a similar absolute increase as with the gold reserve, it would represent a rise of just over 3 per cent. In other words, the same absolute increase represents different *relative* changes. Occasionally it is more important to measure these relative changes than the absolute changes. For example, two firms, A and B, make profits over a period of four years as follows:

Year		1	2	3	4
Profits £000's	A	20	40	80	100
	B	140	160	200	220
Absolute increase £000's ..	A	—	20	40	20
	B	—	20	40	20
Percentage of relative increase each year on previous year	A	—	100	100	25
	B	—	14	25	10

Whereas the first firm has increased its profits fivefold over the same period, an identical yearly rise in the second firm's profits represents much smaller proportionate increases. In such cases where the series extended over a number of years and interest centres on the rate of change in each year, that aspect of the series can be emphasised in diagrammatic form by the use of a semi-logarithmic scale graph as is depicted in Figure 11. Whereas the abscissa is marked off in years as on a conventional graph, the ordinate is scaled off with the logarithms of the values to be plotted.

Figure 11

PRIVATE CAR REGISTRATIONS IN G.B. 1946-56 LOGARITHMIC SCALE.



Based on data from Table 18.

The data on which the graph in Figure 11 is based is given in Table 18 below. For each year the value of the dependent variable, *i.e.*, the number of cars licenced, the logarithm has been written in to facilitate the construction of the graph. Since only

TABLE 18

NUMBER OF PRIVATE CARS WITH CURRENT LICENCES 1946–1955

Year	Number (000's)	Logarithm	Year	Number (000's)	Logarithm
1946	1,770	3·2480	1951	2,380	3·3766
1947	1,944	3·2887	1952	2,508	3·3993
1948	1,961	3·2925	1953	2,762	3·4412
1949	2,131	3·3286	1954	3,100	3·4914
1950	2,258	3·3537	1955	3,526	3·5473
			1956	3,888	3·5784

Source: Annual Abstracts of Statistics. – Based on September quarter registrations.

one scale is logarithmic, it is more correct to refer to such graphs as 'semi-logarithmic' graphs. Certain types of graph are prepared where both scales are logarithmic but they are not relevant here. Furthermore, because of their specialised use, the people who use the 'semi-log' scale graph generally refer to their diagrams as 'log scale' graphs without any risk of confusion. The merit of this graph is that it emphasises the *rate* of change. A constant annual increase in absolute terms, *e.g.*, 200,000 cars a year more every year would, if plotted on a conventional graph, be represented by a straight diagonal line. With the 'log-scale', such a rate of growth would produce a curve which after rising quickly would flatten out. When, as is the case in Figure 11, the curve continues to rise at a constant rate then it signifies that each year the increase is proportionately as great as was the increase in the previous year. It is the *slope* of such a curve which is important; the steeper the rise the more rapid is the expansion in the plotted variable. Note too, that there is no origin marked '0' as with the conventional graph. For this reason it is quite easy to plot a number of such log scale graphs one above the other, each with its own part of the scale. However divergent the series may be, for example, one may be expressed in tens of units per annum and another in millions, this does not create any difficulties. The log-scale reduces them all to a common base and measures the relative change from one year to another in each series, and each line is directly comparable with all the others.

To draw such a graph one can either use specially prepared semi-log-scale paper so that when the actual values are entered on the scale the relative increase between two large absolute values will take up the same space as an identical relative increase between two different and smaller values. Usually, however, the student has no such paper. He can, however, derive its equivalent by setting the log scale of his slide rule against the ordinate and marking off the actual values. This has the same effect as if he had log-scale paper. Finally, the method most generally employed is to extract the logarithms of the numbers and plot them as in Figure 11 above on a natural scale. That is, the difference between log. ·2500 and log. ·3000 is the same scale as that between logarithm ·5000 and ·5500. Whichever method is employed, the same graph should result!

The Histogram

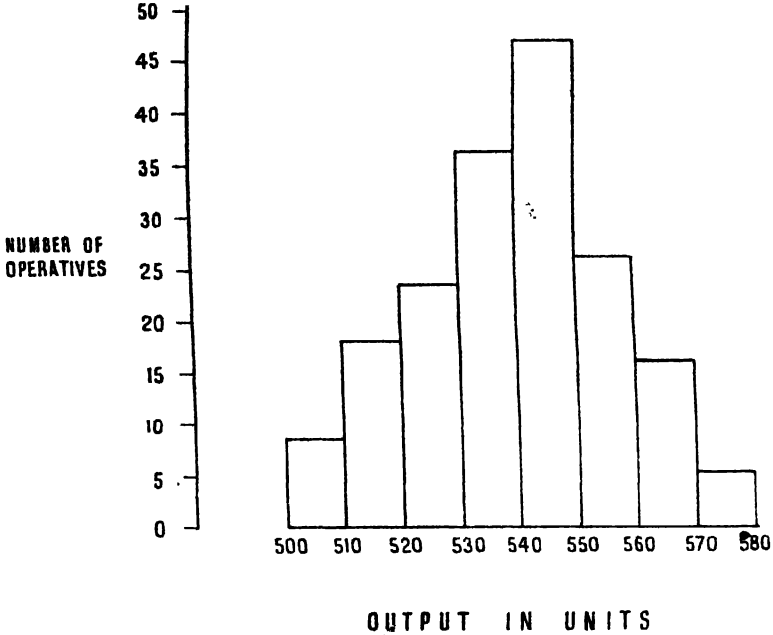
The data given in Table 3, *i.e.*, the outputs of factory operatives, can be plotted, as depicted in Fig. 12. This diagram, comprising what appears to be a series of contiguous rectangles, is known as a *histogram*. This particular diagram is commonly used to depict data given in the form of grouped frequency distributions. For this type of diagram the independent variable is plotted along the base and the frequencies against the vertical. In this example the outputs per class are measured along the horizontal axis and the total frequencies in each class are read off the vertical axis.

The histogram is most easily constructed by drawing vertical lines from the mid-points of each class interval to the height representing the frequencies in successive classes measured off against the ordinate. If flat lines representing the width of each class interval are drawn across the top of each vertical, which then forms the mid-point of each flat bar, and the limits of the horizontal bars joined by perpendiculars to the base, a series of contiguous rectangles is obtained. The aggregate of the frequencies is represented by the total area of the histogram, while the areas of the different rectangles are proportional to the frequencies in the respective classes.

This means that if the class interval is constant, *i.e.*, the same throughout the frequency distribution, the frequencies within each class are indicated by the height of the bar. If, however, as is

Figure 12

HISTOGRAM



Source: Table 3.

sometimes the case the class-interval varies, then the height of the bar must be adjusted. For example, the final class or two often contain relatively few cases so that they are often merged, the class interval of the final group then being twice that of all the other classes. When this class is plotted in a histogram, the length along the base will be twice that of all other classes, *i.e.*, the bar will be twice as wide because the class-interval is twice as large as the others. But it will then have its height reduced by half of the total frequencies in that class so that the area of the bar is proportional to all the others. For the time being the histogram or block diagram as it is sometimes called may be regarded as no more than another method of diagrammatic representation. When the principles of sampling are discussed, it will be seen that the tendency for certain types of frequency distribution to conform to this rather peaked and symmetrical

distribution depicted in Figure 12 is of great value to the statistician.

Continuous and Discrete Variables

Suppose we are classifying households in a given town by reference to the number of children of school age in the household: then we should get a frequency distribution in which the independent variable would take values of 0, 1, 2, 3, 4, etc., according to the number of children in the household. Such measurements are exact, we cannot have less or more than a whole child. The unit is clearly and unequivocally defined for us. Such a variable is termed 'discrete' as is any other variable which can take only certain restricted values, *e.g.*, the distribution of theatre tickets sold during the week classified according to their individual price, *e.g.*, 3s. 6d., 6s., 8s. 6d., etc.; the number of living rooms in a house, and so on. On other occasions, however, the limitations of our measuring rod tend to give approximate values. For example, if we take the temperature of a furnace at one minute intervals, the best readings we get will be rounded to the nearest degree centigrade. If we measure the height or weight of schoolchildren, the recorded heights and weights will be expressed to the nearest unit practicable, *e.g.*, ½″ or 1 pound. Two children may be identical in height and weight, yet the records show one to be ½″ shorter than the other but one pound heavier. The difference is partly explained by the human error in taking these measurements, but it is partly the fault of our height measures and scales which will only given an approximate result, *e.g.*, to the nearest inch or pound, although in practice the approximation is quite adequate for normal purposes. Where a variable can take any value within the range of its observed minimum and maximum values, it is referred to as a *continuous* variable.

The importance of this distinction between the two types of variable is discussed further in the chapter on averages (page 94). It is also relevant in discussing the appropriate form of graph for depicting certain data. Take, for example, the distribution of households by the number of children per household. A line graph with the frequencies plotted against the ordinate and the values 0, 1, 2, etc., along the base would not make sense. It could

be interpreted from the chart that there were (say) 549 households with 1·6 children. A continuous line implies that a reading can be taken from the line at any point. Thus in Figure 9 even if the actual data plotted were half-yearly totals of gold and dollar reserves, a reading between any two dates will still give a sensible if not precise result. For example, if at June 30 the reserves were £800m and by December 31st they totalled £1,000m, it would not be silly to suggest that half-way through that period the size of the reserves was approximately £900m.

As a general rule, line charts and histograms in which the bars are contiguous should be kept solely for plotting the course or distribution of given values of a continuous variable. If the variable is discrete, then column diagrams, such as Figure 6, with a gap between each bar or column are conventionally used. But, in some cases the distribution is such that although the variable is discrete it can be regarded for practical purposes as continuous. The reason for doing this is that if we can assume a variable to be continuous, the statistician can use certain statistical techniques which are extremely useful but which lie beyond our purview. The point can be illustrated by noting the histogram in Figure 12 which relates, strictly speaking, to a discrete variable. After all, a single article produced in a factory can be no more nor less than one unit. But since the differences between successive values of this variable are so small in relation to the value of any single worker's output, *e.g.*, 580 units, we are in the same position as with temperature recordings given to the nearest unit which we stated represented a continuous variable. The same is true of frequency distributions of money sums where the difference between successive recorded values of the variable are small in relation to the individual values themselves.[1]

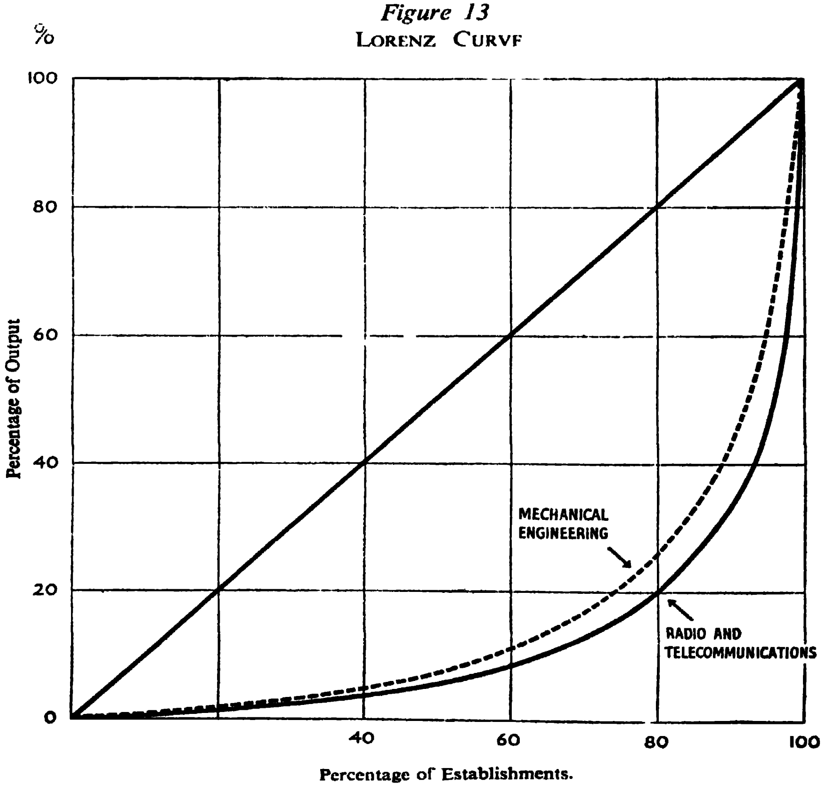
Lorenz Curves

Table 9 on page 47 gave an analysis, among other information, of the firms in the radio and telecommunications industry in 1954 by size – defined by reference to the number of employees – the number of establishments in each size group and their corresponding net output. To ascertain the relative importance of the smaller and larger firms in that industry we could calculate the proportion of total establishments employing above

[1] This point is discussed further on pp. 95-6

a given number of employees, and the corresponding proportion with a smaller labour force per firm. Then for each of these two major groupings we could, by reference to the figures of net output, calculate the corresponding share in the total output. For example, 483 of the 526 establishments listed employ less than 1,000 workers apiece, *i.e.*, 93 per cent of all establishments. This proportion of establishments, however, produces only £58,118,000 worth of output in a total of £138,972,000, *i.e.*, only 42 per cent. In contrast, the firms employing 1,000 workers or more apiece represent only 7 per cent of the *number* of establishments in the industry, but they account for 58 per cent of its net output. Such facts as these for any particular size group of establishment can be derived by graphical methods. The graph in question is known as a Lorenz curve and the curve for these data

Figure 13
LORENZ CURVE



given in Table 9 are portrayed in Figure 13 above. Note that both the X and Y scales are marked off in percentage terms, *i.e.*, 0 to 100. Another curve depicting the distribution of firms and their share of net output in the mechanical engineering industry is also shown in the same graph.[1]

The purpose of such a graph is to illustrate the degree of inequality in the distribution, or the concentration, of output within the industry. Suppose each grouping of firms by size contained 10 per cent of the number of firms and each group accounted for 10 per cent of the net output of the industry. In such a case, the resultant curve would be the diagonal straight line on the graph which illustrates perfectly equal distribution. As was apparent from the figures quoted above, in the radio industry the concentration of output in the few large firms is quite marked and this is brought out by the curvature of the line. The flatter the curve, *i.e.*, the more closely it approaches the diagonal straight line, the more equal the distribution. Thus, the second curve for the mechanical engineering industry reveals a slightly less uneven distribution; or in other words, the concentration of output in the bigger firms is not quite as marked as it is in the radio industry.

The student could draw a similar curve for the data in Table 6 which gives the distribution of personal incomes in 1955-56. For each figure of incomes and total income in the second and third columns, calculate the percentage it forms of the total. For example, the first figure in the second column is 2,075 which is roughly 10 per cent of all incomes. Their share of total incomes is only £443m. out of £12,374m., or about 3·5 per cent. When each percentage in both columns has been calculated, cumulate the successive percentage figures and plot those cumulated figures on the graph with both scales marked off in percentages, so that each pair of percentages from columns 2 and 3 will provide one point on the graph.[2] There are nine pairs of figures, so there will be nine points which when joined will form the curve. Note that the curve starts at zero and finishes at 100 on both scales. Thus the first point will be located 10 per cent up the income axis and 3·5 per cent along the other.

[1] Based on data in Table 3 Volume 4, Industry 1 of the Report on the Census of Production for 1954.

[2] The student who has forgotten the process of cumulating such distributions should refer to Table 4 on page 41.

Summary

The type of diagram to be used for depicting given data does not usually pose serious problems. As with tabulation, diagrammatic representation of statistical data is largely a matter of commonsense coupled with a few rather obvious rules. Since experience shows that these rules tend, all too often, to be overlooked, it may be helpful to set them out below. The charts and graphs in the previous pages and elsewhere in the book can then be assessed in the light of these rules.

1. The title should be brief but self-explanatory.

2. The source of the data plotted should be given so that the reader may consult them for himself.

3. The axes of the graph or chart must be clearly labelled so that the quantities and values to which they refer are immediately apparent.

4. Except for logarithmic graphs, always show the origin. Whenever the values are such that the points plotted will normally lie a considerable distance from the origin and lead to a compression of the scale, then the graph may be 'broken' or 'torn' across the vertical axis and the relevant values marked off along virtually the entire length of the ordinate.

5. Always bear in mind that graphs are meant to simplify the picture; a detailed or overloaded chart or graph defeats its own object.

Finally, it should not be forgotten that the addition of a diagram or graph among the published data adds nothing to what is already available. Its sole *raison d'être* is that it drives home more effectively than the tabulated data the main findings of the enquiry. If the diagram does not meet this requirement, then it is better omitted. Generally speaking, diagrams are useful aids to

comprehension. The reader interested in charts will find a varied collection in any issue of *Economic Trends* published monthly by H.M. Stationery Office.

CHAPTER VI

AVERAGES: TYPES AND FUNCTIONS

The Function of Averages

Few people can assimilate a mass of detailed information expressed in numerical form, even when it has been substantially reduced by tabulation. It is helpful, therefore, if instead of merely tabulating the information derived from a specific enquiry and depicting it in graphs or diagrams, it can be expressed in more abbreviated numerical form, yet in such a way that the salient features of the tables are clearly brought out.

For instance, in the case of the firm owning the plant with 180 employees whose outputs were given in Table 1 it may be assumed that this firm controls several such plants of varying size in different parts of the country. The management would be anxious to compare the outputs of the operatives in the various plants. If conditions in each plant are similar, the results should closely correspond. If there are serious discrepancies in their respective production levels, then some explanation must be found. It would be a tedious process comparing all the individual outputs in every plant, finding out the lowest output, the highest, and the most frequent, by such tables and graphs as we have so far employed for illustrative purposes. If all the significant features of the data relating to each plant can be brought out by one or two figures, their comparison is a far simpler task than the detailed scrutiny of the data suggested above. These 'summary figures' may for the moment be described simply as 'averages', illustrated by the following three examples:

1. If, for example, it is stated that the average weekly output of an operative in Plant 1 is 539 units and in Plant 2 with identical working conditions it is 519, such information warrants investigation.
2. Further, if more operatives each produce 539 units per week than any other output in Plant 1, and the corresponding most frequent output in Plant 2 is 515, the apparent conclusion is that the operatives in Plant 1 are for some reason generally more productive.

3. If the individual outputs of the operatives in both plants are ranged separately in two arrays, *i.e.*, in order of magnitude as in Table 1, it may appear that the middle worker in the array for Plant 1, *i.e.* the 91st out of 180, has a weekly output of 540 units, while in Plant 2, with an equal number of operatives, 120 have a smaller weekly output than this.[1] The management confronted with this information would invariably seek an explanation.

Given the facts above, together with the ranges of the two distributions, it would be possible for anyone conversant with statistical methods to estimate with reasonable accuracy the general level of productivity in each plant, and even depict distribution of outputs as a frequency curve sufficiently accurately to bring out the same essential features as a graph of the complete data. It is because 'averages' summarise the salient features of most data so usefully that they are so widely employed in statistics. In fact, statistics has been described as 'the science of averages', although this is a little misleading in so far as averages form only a part of the techniques employed, particularly in the later stages of a statistical enquiry.

The three specific comparisons made above have now to be considered separately and in detail.

The Arithmetic Average or Mean

'The "average" output of the operatives in Plant 1 was 539 units.' Most people are acquainted with the use of the term 'average' in this context. Thus in cricket, when a batsman has 'averaged' 50·0 runs per innings, no one assumes that exactly 50 runs have been scored in each innings or possibly in any innings; but if the total runs scored are divided by the number of completed innings the result will represent the batting 'average'. Assuming the above batsman completed 30 innings it is clear that his aggregate is 1,500 runs.

Using the data in Table 1, by aggregating the individual outputs and dividing the total by 180, an average output per operative of 539 units is obtained. Reference to the detailed array reveals that only five workers actually produced this output, and with only this figure as a guide, a somewhat limited picture of the situation would be obtained. This fundamental weakness in

[1] To be exact, the *middle* operative in a series of 180 'lies between the 90th and 91st'. This point is discussed later in the chapter, it does not affect the present argument.

this type of average, or arithmetic 'mean' as it is known in statistics, is even more clearly demonstrated by the following example. A prospective investor is informed that three companies, X, Y and Z, have during the past six years each averaged a net profit of £6,250 each. It is assumed for the purpose of this example that the annual profit figures have been adjusted to a common basis in order to eliminate non-recurring or capital items and are therefore comparable. It may be assumed that the actual figures for each company over the period are as follows:

Year	X Co. Ltd.	Y Co. Ltd.	Z Co. Ltd.
	£	£	£
1948	1,500	8,800	12,000
1949	4,000	7,200	4,000
1950	7,000	6,600	*Loss* 2,000
1951	8,000	6,000	8,000
1952	8,000	5,900	15,000
1953	9,000	3,000	500
	6)37,500	6)37,500	6)37,500
Average Annual Profit ..	6,250	6,250	6,250

It will be self-evident that confronted with this more detailed information the investor would promptly forget all about the 'average' profits and completely revise his first impression based on the original statement of equal average profits that the investments are equally attractive. Thus the bare 'mean' may give quite a misleading impression of any series or distribution as it provides no indication of the variation between the actual values within the distribution.

Calculation of the Mean

In practice, however, the Arithmetic Mean is not always quite so simple to compute, as is illustrated by the following example. In a certain works, the works staff comprises 100 skilled men, 200 semi-skilled operatives and 50 unskilled men, all of whom are paid on a time basis at £15, £10 and £8 per week respectively. The 'average' or 'mean' wage paid in the works is *not* £11 per

week computed as follows: $\frac{£15+10+8}{3}=\frac{£33}{3}$, *i.e.*, £11 per week. The inaccuracy of this result may be easily proved. The total sum required to pay the weekly wages of the factory is £3,900, *i.e.* (£15 × 100, £10 × 200, £8 × 50). The total yielded by multiplying the first mean of £11 by £350 workers is £3,850, or £50 short. The correct mean wage is £3,900 divided by 350, *i.e.* £11 2s. 10d.

This type of average is sometimes described as a 'weighted' mean, *i.e.*, the separate values or items within the series are each multiplied by the frequency with which each item or value appears. In the preceding example, the weights were the number of employees within each group, 100, 200 and 50 respectively. Such a computation is required when a compound made up of several constituents has to be priced for the purpose of cost accounts or final stock valuation. Thus, if A, B, C and D are four chemicals costing £15, £12, £8 and £5 per cwt. respectively, and are contained in a given compound in the ratio of 1, 2, 3 and 4 parts respectively, the resultant compound must be priced out at £ (1 × 15) + (2 × 12) + (3 × 8) + (4 × 5) divided by 10, equalling £8 6s. 0d. per cwt. In the correct statistical sense of the term these figures (numbers of workers or cwt.) are *not* weights – they are simply the frequencies of each value of the independent variable. Unfortunately, the distinction is not always clearly made and all too often the frequency of a single value or group of values is termed its 'weight'.[1]

The calculation of the Mean from a *grouped* frequency distribution is different from that employed for the simple series or frequency distribution above. Where the data have been grouped, the exact frequency with which each value of the independent variable occurs in the distribution is unknown. Our knowledge is limited to the fact that, within successive class limits of the independent variable, a certain number of frequencies occur. The procedure for calculating the Mean in such cases is illustrated in Table 19.

For the purpose of 'averaging', the mid-point of the class-interval is selected. This arbitrary procedure is justified on the score that if the number of frequencies is large, the frequencies within each class will probably be spread evenly over the range of

[1] The main use of weighting is discussed in the Chapter on Index Numbers.

TABLE 19

REJECTS PER OPERATIVE IN PLANT 4 DURING 4-WEEK PERIOD ENDED 8TH NOVEMBER 1958

(1) No. of Rejects	(2) Mid-point	(3) No. of Operatives	(4) Products of cols. 2 × 3
21-25 ..	23	6	138
26-30 ..	28	17	476
31-35 ..	33	22	726
36-40 ..	38	34	1,292
41-45 ..	43	20	860
46-50 ..	48	12	576
51-55 ..	53	5	265
		116	4,333

Average rejects per operative $\frac{4,333}{116}$ 37 to nearest unit.

the class-interval, *i.e.*, there will be as many items below the mid-point as above it.[1] Thus by using the mid-points for calculating the average the same result would be obtained as that given by aggregating the products of the individual values and their respective frequencies. It should be noted that the procedure of multiplying the mid-points of the classes within a grouped frequency distribution by the number of items within the respective classes does not actually provide the Mean as such. It produces the Mean of all the mid-points of the classes 'weighted' by the frequencies within each class of the distribution. Since, as stated above, the assumption is made that the Mean of all the values within each class is equal to the mid-point of that class, the use of the mid-points in order to obtain the Mean of the distribution is permissible. Generally, the smaller the class-interval and the larger the number of frequencies in each class, the more likely it is that the 'mid-point' average will correspond

[1] The same arithmetic result will, of course, be obtained if the majority of the frequencies were concentrated on the mid-point of the group as in (2), and the remainder of the frequencies spread equally over the other four values in the group, *i.e.*, two on each side of the mid-point. Using hypothetical figures we get:

(1) Value in units	f	$f.x.$	(2) Value in units	f	$f.x.$
1	16	16	1	5	5
2	16	32	2	10	20
3	16	48	3	50	150
4	16	64	4	10	40
5	16	80	5	5	25
	80	240		80	240

Average = 3 Average = 3

The student should note that these results arise because the frequencies are distributed symmetrically about the mean.

to the average calculated exactly *i.e.*, if the data given as an ungrouped frequency distribution in Table 2 were employed.

When preparing the grouped frequency distribution it will be seen whether all the items are dispersed evenly throughout the range of the independent variable. If this is the case, the classes may be taken at the most convenient intervals, *e.g.*, multiples of 5 or 10 units as with the data in Table 1 (p. 39). But where, as is frequently the case, there are irregular concentrations at intervals throughout the range of the independent variable, the obvious class-limits may not be suitable and it will be necessary, as was explained on p. 42, to revise the class intervals in such a way that within each class the mean value will be found around the mid-point. Clearly this is an ideal seldom attainable in practice, but it is important to remember it when a simple frequency distribution such as is given in Table 2 is converted into the grouped distribution shown in Table 3.

Determination of the Mid-point

Apart from the questions of selecting a suitable number of classes for the grouping of any frequency distribution, and the size of the class interval, care must be taken to ensure that no uncertainty can arise in allocating any particular value to its appropriate class. The most common slip is to state the class interval as follows: £10-20, £20-30, £30-40 and so on throughout the range of the independent variable. If, after its compilation, a grouped frequency distribution in this form were to be examined, three alternatives concerning the disposition of those units which are multiples of £10, *i.e.*, £20, £40, etc., spring to mind. The person responsible for the classification may have had no system at all, sometimes the item was put in the lower class, sometimes in the upper class; *e.g.*, £30 could have been put in £20-30 or £30-40. The second course would have been to place them systematically in the upper class, *i.e.*, assume that the upper limit of the preceding class was read as 'under £30'. The third alternative would be to assume that the lower limit of the class £30-40 meant all items *over* £30 to be included. The value of any calculations performed on such dubious tables would be problematical to say the least. In the chapter on Tabulation, the need for accurate classification was stressed, and the above example should emphasise the reasons.

The classes quoted above can be written in several ways, and although the differences are not important one method may be more suitable than another for a particular distribution. Thus:

(1)	(2)	(3)	(4)
Under £10	£0— Signifying	—£10 Signifying	0 — £9
£10 and under £20	£10— up to but	—£20 up to and	£10 — £19
£20 „ „ £30	£20— not inclu-	—£30 including	£20 — £29
	ding £10,	£10, or	
	£20 etc.	£20	

The first is clear enough and can be used for values quoted to the nearest penny. A frequently used alternative to the classification in (1) is given in (2); they are the same. The third example differs from (2) since a value of exactly £10 will fall in the second class in the first example but in the first class in (3). The last grouping is based on the assumption that all the items are given to the nearest pound. The conventional methods for deriving the mid-points of the classes in a distribution are as follows. If the variable is *discrete*, and the classification written as follows: 1 — 5, 6 — 10, 11 — 15, and so on, the mid-points are clearly 3, 8 and 13 respectively. The method may be summarised by stating that the limits of each class are aggregated and their sum halved, *e.g.*, $\frac{6+10}{2} = 8$.

Continuous variables are slightly more difficult, since much depends on the correct demarcation of the class limits. If the classes are written as in example (1) above, 'Under £10', and so on, the class limits will depend on the way in which the individual values in the distribution have been expressed. For example, if all are expressed to the nearest penny, the upper limit of the first class is £9 19s. 11d.; the limits for the second class are £10 and £19 19s. 11d., and so on. Strictly speaking, money values should be treated as a discrete variable, but as for example in the above illustration, the smallest unit of one penny is so very small in relation to the individual values, the error introduced by treating the values as continuous may be ignored. Normal practice with *continuous* variables is to derive the mid-point by halving the sum of the lower limits of two successive classes. Applying this rule to example (3), the mid-point for the second class would be £15. If the values in a continuous distribution are written to

the nearest decimal place, the classes could then be written — 10·0, — 20·0, — 30·0, — 40·0, etc., with mid-points of 5·0, 15·0 and 25·0.

When in doubt as to the limits of the class interval the student should consider firstly the unit of measurement and secondly how the individual values have been defined.[1] For example, if no payment is less than a multiple of a pound then the series is discrete, since the difference between the upper limit of a class interval and the lower limit of the next must be one pound. Such a classification would then read as in example (4). If, however, the values have been rounded to the nearest pound, then clearly a value of £10 in the distribution could represent any value ranging from £9 10s. 0d. to £10 9s. 11d. Such a distribution should be treated as a continuous series and the classification should be similar to that given in any of the first three examples. It should be remembered that the grouping of a distribution and the use of mid-points for calculating averages by themselves give rise to possible error in the average. The choice of the class-interval should be as accurate as is compatible with such considerations.

The Short-cut Method

Using the mid-points of successive classes in a frequency distribution to compute the Mean, the volume of arithmetic calculation can be reduced in the following ways:

TABLE 20

INDIVIDUAL OUTPUTS OF 180 FEMALE OPERATIVES AT PLANT 1 IN THE WEEK ENDING 8TH NOVEMBER 1958

Output in Units (1)	Mid-points (2)	Mid-points less 504·5 (3)	No. of Operatives (4)	Products (3) × (4)
500 to 509 ..	504·5	Nil	8	Nil
510 to 519 ..	514·5	10	18	180
520 to 529 ..	524·5	20	23	460
530 to 539 ..	534·5	30	37	1,110
540 to 549 ..	545·5	40	47	1,880
550 to 559 ..	555·5	50	26	1,300
560 to 569 ..	565·5	60	16	960
570 to 579 ..	575·5	70	5	350
			180	6,240

Mean output per operative = 504·5 + 34·7 = 539 to nearest unit.

[1] And if he has forgotten the difference between discrete and continuous variables he should re-read p. 83.

In this example, by using the mid-points and subtracting the figure of 504·5, which is common to every mid-point, the arithmetic involved is reduced to very simple proportions.

Such a simple example involving easily manageable figures does not arise very frequently, and a more usual method of computing the arithmetic mean is given in Table 21 below. This method is based upon the simple rule of algebra that the sum of the individual differences between a series of numbers and their mean is always equal to zero. Take for example the following distribution: 4, 7, 9, 10, 15, 17 and 22. Their aggregate is 84 and the average of the seven figures comprising that total is therefore 12. From this figure, *i.e.* their mean, subtract each of the figures in the distribution in turn. The following result is obtained:—8, —5, —3, —2, 3, 5, 10. When aggregated the differences are equal to zero. The reader may check the rule by testing any selection of values he cares to make. Given this rule the accuracy of an estimated mean may be tested quite simply. Suppose that for the above series we had guessed that the true mean equalled 10. The differences, or as they are usually termed, the *deviations* from the mean, would then be: —6, —3, —1, 0, 5, 7, 12 and their total is 14. Clearly then, if the rule is valid the estimate of the mean is wrong. But if the 'error' is apportioned, *i.e.* 14 units, between the seven constituent numbers, the 'average deviation' is 2 and if this is then added to the estimated value of the mean, *i.e.* 10, we arrive at the correct value of the mean of the distribution, *i.e.* 12. The student should amuse himself setting out short series of figures and proving to himself the validity of the rule. As will be seen, it forms the basis of many calculations in statistics and it is used to calculate the mean of the distribution shown on p. 98.

The foregoing principle is illustrated in Table 21. The successive stages in the calculation are as follows:

1. Select as the assumed mean the mid-point of the class which contains a high proportion of the units and is nearly central. Since wages are paid to the nearest penny, the true limits of the first class are £4 10s. — £4 14s. 11d. and the variable is, strictly speaking, discrete. But, for all practical purposes, this variable may be treated as continuous since one penny is so small a unit and the mid-points are derived by adding together the upper limits of successive classes and halving them.

TABLE 21

EARNINGS OF 1,783 FEMALE EMPLOYEES OF THE XYZ MANUFACTURING CO., LTD. IN THE WEEK ENDING 8TH NOVEMBER 1958

Weekly Earnings (1)	Frequency (2)	Mid-points (3)	Deviations from assumed Mean in Class Intervals (4)	Products of cols. ~~3~~× 4: Negative (5)	Products of cols. ~~3~~× 4: Positive
£ s. d. £ s. d.					
4 10 0 but under 4 15 0	64	4·12·6	— 4	— 256	
4 15 0 " " 5 0 0	126	4·17·6	— 3	— 378	
5 0 0 " " 5 5 0	224	5· 2·6	— 2	— 448	
5 5 0 " " 5 10 0	379	5· 7·6	— 1	— 379	
5 10 0 " " 5 15 0	474	5·12·6	0		
				— 1,461	
5 15 0 " " 6 0 0	227	5·17·6	1		227
6 0 0 " " 6 5 0	108	6· 2·6	2		216
6 5 0 " " 6 10 0	74	6· 7·6	3		222
6 10 0 " " 6 15 0	31	6·12·6	4		124
6 15 0 " " 7 0 0	43	6·17·6	5		215
7 0 0 " " 7 5 0	19	7· 2·6	6		114
7 5 0 " " 7 10 0	14	7· 7·6	7		98
	1,783				1,216

Assumed mean = mid-point of the class £5 10s. 0d. but under £5 15 s. 0d. = £5 12s. 6d.

Sum of deviations from assumed mean = — 245 i.e., — 1,461 + 1,216

Average deviation in class intervals = $\frac{-245}{1{,}783} = \cdot137$

Correct arithmetic mean £5 12s. 6d. — ·137 of the class-interval
= £5 12s. 6d. — ·137 × 5s. 0d.
= £5 12s. 6d. — 8d. to nearest 1d.
Arithmetic mean = £5 11s. 10d. to nearest 1d.

2. In the column headed 'deviations from assumed mean in class-intervals', enter 0 against the class whose mid-point is to be used as the assumed mean, *i.e.*, £5 10s. 0d.—£5 15s. 0d. Against the mid-points on either side of this latter class enter 1, against the next above and below those mid-points enter 2 and so forth. Where the mid-point is smaller than the selected mid-point representing the assumed mean, the deviation will be negative; thus the upper part of Col. (4) before the mid-point marked 0 will contain all the negative deviations. The

reverse applies to the lower part where the mid-points are greater in magnitude than the assumed mean, *i.e.*, the deviations are positive. Before inserting the figures, the student should note whether all the class-intervals are of equal size; *i.e.* as in Table 21.

3. Each deviation is multiplied by its respective frequency, *i.e.*, the frequencies in each class; the negative quantities being kept apart from the positive products to avoid confusion. Both the negative and positive products are then aggregated separately and the balance obtained.
4. The balance, negative or positive, is divided by the sum of the frequencies. The result in this example is a *fraction of the class-interval*, not of a single unit. In other words the result is expressed in 5s. units, since it was derived by dividing the net sum of the products of —245 by the total frequencies of 1,783. Reference to columns (3) and (4) will reveal that the deviations are measured in units of 5s; thus 2 deviations equal 10s. as is apparent by subtracting the mid-point of the class £5 10s. 0d. — £5 15s. 0d. from that of the class £5 0s. 0d. — £5 5s. 0d. It is important therefore that the quotient of —·138 is converted into shillings before it is subtracted from the assumed mean, from which the deviations were measured. Unfortunately many students forget this small point in their examinations. The result gives the true arithmetic mean of the frequency distribution.

The figures entered under the heading of 'deviations from the assumed mean', might have been multiples of 1s. instead of 5s., or, for that matter, of any unit always provided the difference between the negative and positive totals expressed in terms of those units is finally converted to the original unit of measurement. If the deviations had been measured in actual shillings instead of multiples of 5s., then the figures in column (4) would have read 0, 5, 10, 15 and so on instead of 0, 1, 2, 3 etc. Equally, the '0' *i.e.* the mid-point assumed to be the mean from which all the deviations are measured, may be placed anywhere in the series, but it simplifies the calculation if put against the class limits between which the largest number of frequencies occur.

It will be realised that there is no need to work in class-intervals, although this is usually the most convenient method

when all the class-intervals are equal in size. If, however, they vary, the mid-point method of deriving the Mean may still be used. With varying class-limits the differences will be in multiples (sometimes fractions) of the class-interval 'unit'. If, for example, there had been another class at the lower end of Table 21, say £7 10s. 0d. to £8 5s. 0d. then the mid-point of this class interval is £7 17s. 6d. Had this been the case the deviation from the mean allowed for that class would be 9 and not 8, since the difference between the mid-point of this class and that immediately preceding it is equal to 10s.; twice as much as the unit of 5s. in which previous deviations have been measured. A little extra care is necessary, therefore, in writing in the values in the 'deviations' column. Where 'open-end' classes are involved, *e.g.* '£7 10s. and over', the difficulty is still greater. If the open-end was necessary there is every reason to assume that one or more of the frequencies did not fall within the limits of any normal class, *i.e.* there is (or are) extreme item(s) which would affect markedly the value of the Mean. The usual assumption in the absence of further information is to assume the limits of that class are identical with the others and select a mid-point accordingly. It is probable that the use of this arbitrary mid-point will tend to under-estimate the true Mean of the distribution. Since extreme or unrepresentative items distort the Mean this compromise avoids that danger, but the method is still unsatisfactory. The only guide is knowledge of the data being handled. A worked example illustrating the problem of open-end classes and varying class-intervals in the same distribution is given at the end of the next chapter.

Apart from its relative simplicity of computation, the arithmetic Mean has other advantages. The statistician considers it a useful measure since it is itself the result of an arithmetic process and therefore lends itself to further mathematical treatment. Insofar as the Mean takes into account every item in the series or distribution, it generally provides a reasonably accurate summary of the data, hence its popularity with the lay public who use the term 'average man' to refer to the representative of the majority of the male members of the community. On the other hand this advantage also lies at the root of its outstanding weakness; by including every item in the series the presence of extreme or single non-representative items may, especially in a

short series, so seriously distort the Mean that it no longer provides an accurate indication of the nature of the data. Thus, if seven directors receive annual fees of £100, £200, £200, £250, £250, £300 and £1,500 respectively, the Mean is £400. In fact, this particular amount is not received by any director, is in excess of what six out of the seven receive and provides no indication whatsoever of the nature of the series on account of the extreme item of £1,500. This weakness occasionally provides one of the reasons for needing other measures which will amplify and even replace the Mean, although the latter remains one of the most frequently employed measures in statistical work.

Formulae for the Mean

Reference to the majority of books on statistics will reveal a formidable array of mathematical symbols dealing with the calculation of the Mean, which convey little to the non-mathematical reader, and may even confuse the issue for him.

Most of the symbols employed are merely 'shorthand' or abbreviations of simple procedures which would be cumbersome if expressed in simple English.

Thus:

1. The Arithmetic Mean of a *simple* series, *e.g.*, 2, 6, 9, 12, 15, is written $X = \frac{\Sigma x}{\Sigma f}$ to represent

$$\text{A.M.} = \frac{(2 + 6 + 9 + 12 + 15)}{5}$$

X = the arithmetic mean.
where Σ (termed large sigma) = the sum of.
x = the individual items.
Σf = total frequencies

By using Σ the need for the following notation is avoided:

$$\frac{x_1 + x_2 + x_3 + x_4 + \ldots x_N}{N}$$

where x_1, x_2, $x_3 \ldots x_4$ refer to the individual values of the variable and N = the total frequency.

2. For frequency distributions, such as the wages or chemical compound examples on page 92, the formula may be written:

$$\frac{f_1x_1 + f_2x_2 + f_3x_3 \ldots + f_nx_n}{f_1 + f_2 + f_3 \cdots + f_n} \text{ i.e., } \frac{(15 \times 1) + (12 \times 2) + (8 \times 3) + (5 \times 4)}{(1 + 2 + 3 + 4)}$$

which is normally abbreviated to $\frac{\Sigma fx}{\Sigma f}$, the letter f representing the frequencies and Σf, it should be noted, is the same as N.

3. The A.M. of a *grouped* frequency distribution is written $\bar{X} = \frac{\Sigma fx}{\Sigma f}$, where $f =$ the number of observations in each class of the distribution. Note that x in this case represents the mid-points of the classes.

4. When the A.M. is derived from a frequency distribution by using the *deviations from an assumed Mean*, the formula is written $\bar{X} = X' + \frac{\Sigma fd'}{\Sigma f}$, where $d' =$ the deviations from the assumed Mean written as X'. (Note: if the deviations are expressed in class-intervals, they must be converted into the original unit values. Thus $\frac{\Sigma fd}{\Sigma f} \times i$). This formula applies to the example in Table 21. The student reader may commit these formulae to memory in case they should appear on an examination paper, or more probably, in another text. For practical purposes they are unnecessary at this stage, it is the method, not the formula, which should be learnt.

The Median

The nature of the third average employed in describing statistical data was indicated in the passage on page 90. 'If the individual outputs of the operatives are ranged in order of magnitude, *i.e.*, an array, the central figure has a value of 540 units'. The Median divides the distribution into two equal parts, in other words, it is the value which divides a distribution so that an equal number of values lie on either side of it. In this example the one half contains the better operatives and the other the less productive. This is a particularly useful 'average' for distributions which are very markedly non-symmetrical.

Median of Ungrouped Data

In contrast to the Mean the task of finding the Median is sometimes extremely simple. All that is necessary is to arrange

the individual items in order of magnitude; the middle item is then the Median. Thus in the following series, 2, 3, 4, 5, 6, 7, 8. 9 and 10 the Median is 6, *i.e.*, the fifth figure, with four figures on either side of it. Such is the procedure when the data is given in an array, more usually it is described as *ungrouped* data. It may be located by the formula $\frac{N+1}{2}$, where $N=$ the number of items in the series. When the number of items, *i.e.* N, is odd, then the Median is an actual value with the remainder of the series in two equal parts on either side of it. If N is even, then the Median is a derived figure, usually half the sum of the two middle values. If these are the same, as they often are, then the Median of an even series will also be an actual value.

Median of Grouped Data

More frequently, however, it is necessary to select the Median from grouped data, *i.e.*, a frequency table where the original data has already been condensed into classes. In this case the ranging of the data has already been effected since the classes will clearly be in order. The normal method is to add the class frequencies together cumulatively, as has been done in the example below and divide the total frequencies into two halves.

TABLE 22

EARNINGS OF 1,783 FEMALE EMPLOYEES OF THE XYZ MANUFACTURING CO., LTD. IN THE WEEK ENDING 8TH NOVEMBER 1958

Earnings	No. earning Wages shown opposite	Cumulative Total	
£ s. d. £ s. d.			
4 10 0 and under 4 15 0.. ..	64	64	
4 15 0 ,, ,, 5 0 0.. ..	126	190	(64 + 126)
5 0 0 ,, ,, 5 5 0.. ..	224	414	(190 + 224)
5 5 0 ,, ,, 5 10 0.. ..	379	793	(414 + 379)
5 10 0 ,, ,, 5 15 0.. ..	474	1,267	(793 + 474)
5 15 0 ,, ,, 6 0 0.. ..	227	1,494	etc.
6 0 0 ,, ,, 6 5 0.. ..	108	1,602	
6 5 0 ,, ,, 6 10 0.. ..	74	1,676	
6 10 0 ,, ,, 6 15 0.. ..	31	1,707	
6 15 0 ,, ,, 7 0 0.. ..	43	1,750	
7 0 0 ,, ,, 7 5 0.. ..	19	1,769	
7 5 0 ,, ,, 7 10 0.. ..	14	1,783	
	1,783		

The formula for deriving its position from *grouped* data, *i.e.*, as in a grouped frequency distribution, is $\frac{N}{2}$. Thus, in the following example, it is located at the mid-point between the two middle items of the series, *e.g.*, 1,783 items in series – the Median value lies between the 891st and 892nd item. The rest of the calculation is given in full below:

Median: $\frac{1{,}783}{2} = 891\frac{1}{2}$th item.

The Median is located between 793 and 1,267, *i.e.*, among the 474 individuals receiving at least £5 10s. 0d. but under £5 15s. 0d. Therefore the median wage is greater than £5 10s. 0d. but below £5 15s. 0d.

$891\frac{1}{2} - 793 = 98\frac{1}{2}$, thus the Median is the $98\frac{1}{2}$th of the 474 items ranged in order of size, these 474 values ranging from £5 10s. 0d. to £5 14s. 11d.

£5 10s. 0d. + $\left(\frac{98\frac{1}{2}}{474} \times 5\text{s.}\right)$ = £5 10s. 0d. + 1s. 0·3d. or 1s. 0d. to nearest penny, *i.e.*, £5 11s. 0d. = Median wage.

Alternatively the calculation may be carried out by assuming that £5 15s. 0d. is greater than the median wage, and 375 (474 – 99) employees in that class earn more than the median wage.

Median wage = £5 15s. 0d. – $\left(\frac{375}{474} \times 5\text{s.}\right)$ calculated to nearest penny = £5 11s. 0d.

It will be noticed that the same assumption has been made in the computation of the Median as was made in determining the Arithmetic Mean, *i.e.*, that the items falling within any given class are ranged evenly throughout, and as before, the validity of this assumption will determine the accuracy of the result.[1] This is justified with a continuous series with a large number of classes. If the variable is discrete and the class-interval large, the Median may be little better than an approximation, and the result is best given to a round number. The outstanding advantage of the Median resides in the fact that it is not affected by extreme items, as is the Mean. Thus if seven salesmen take £700, £750, £780, £800, £830, £870 and £1,600 respectively, the Median value is £800, which gives a fair indication of the typical salesman's results; the Mean on the other hand is over £900 and quite unrepresentative. The Median value often corresponds to a definite item in the distribution; the Mean seldom. A further important advantage of the Median is that

[1] This assumption need only be valid for the class containing the Median – since the Median is unaffected by the values in any other class.

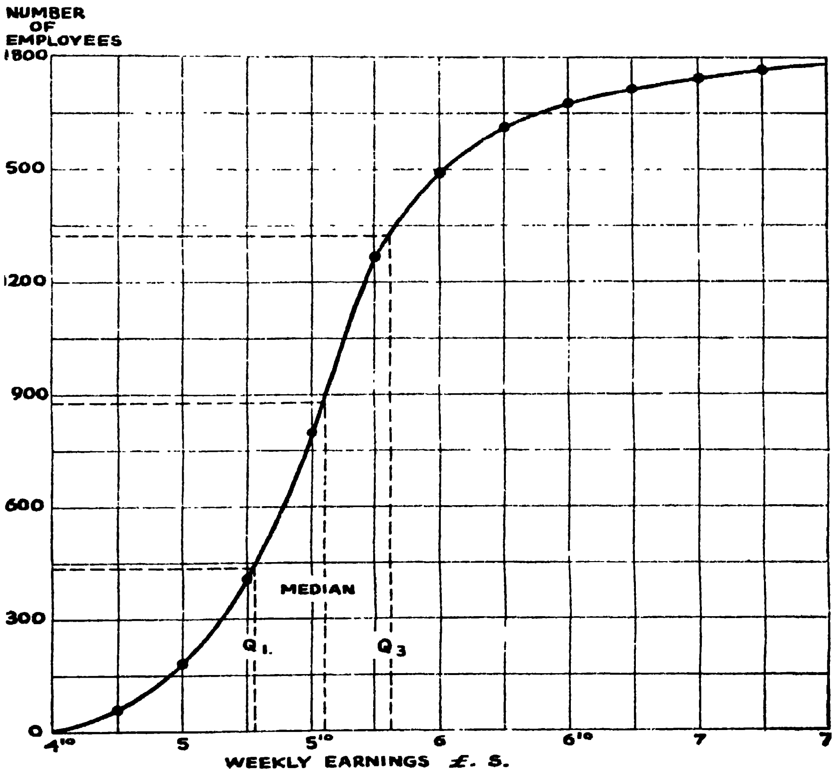
it can be located in a grouped distribution in which the first and last classes are open – ended and the lower and upper units are not available so that it is virtually impossible to compute the Mean with any degree of accuracy.

Median by Interpolation

The Median can also be interpolated approximately from the ogive, *i.e.*, the cumulative frequency distribution plotted on a graph as shown below. This is true whether the ogive is drawn

Figure 14

DEVIATION OF MEDIAN AND QUARTILES FROM A CUMULATIVE FREQUENCY CURVE.



Based on data from Table 22.

by cumulating the series upwards or downwards as was explained on p. 41. Care is required when drawing this curve that the data are correctly plotted. Thus the curve starts at zero frequencies and rises continuously. The first point plotted against the vertical axis, in this case 64 (see Table 22, p. 103) will be above the upper limit of the class £4^{10}—£4^{15}, the next figure 126 against the upper limit of the next class, *i.e.* £4^{15}—5. Thus, when the frequencies of any given output are read from the curve, they will be interpreted as '190 employees below £5, 64 below £4^{15}, etc. The student should not plot the frequencies over the mid-points of the class-intervals as is done with the histogram, otherwise he will read off the wrong results, *i.e.* the values of the independent variable will be too low. Having drawn the ogive, all that is necessary is to find the mid-point on the scale representing the frequencies, which in Fig. 14 is the vertical axis and from that point draw a line parallel to the horizontal axis until it intersects the ogive. The value of the Median will then be read off against the scale along the horizontal x axis directly below the point of intersection. The reader can compare the values derived from the graph with those calculated from the data, the Median on p. 104 and the quartiles below. Since both methods of deriving these values are at best approximate it is not surprising that there are slight differences between the results.

Quartiles and Deciles

In precisely the same way, it is possible to locate the *quartiles* and *deciles*, values of which are sometimes useful in describing a distribution. As the name of the former suggests, the *quartiles* divide the series into four equal parts, *i.e.*, they perform for each equal part of the series on either side of the Median what the Median has done for the whole series. The *deciles*, less frequently employed in practical work than the quartiles, divide the series into 10 equal parts. The method of computing the quartiles of *grouped* data is the same as with the Median, except that instead of $\frac{N}{2}$, the denominator is 4, *i.e.*, $\frac{N}{4}$. The two quartiles in any distribution are known as the *lower* and *upper* quartiles, the former indicating the smaller value and obtained by $\frac{N}{4}$, and the latter, the higher value in the position $\frac{(3N)}{4}$. The lower quartile is

usually written as Q_1, the upper quartile as Q_3. The calculations for the deciles are similar, the denominator being 10, thus the fourth decile is the observation which is $\frac{4N}{10}$ from the lower end of the range.[1]

Working on the data given in Table 22, the following results are obtained:

$Q_1 = \frac{1{,}783}{4} = 446$ to nearest unit, *i.e.*, $Q_1 =$ value of 446th item.

$Q_3 = \frac{3 \times 1{,}783}{4} = 1{,}337$ to nearest unit, *i.e.*, $Q_3 =$ value of 1,337th item.

The 446th item lies in the class £5 5s. 0d. — £5 10s. 0d. containing 379 items, which are assumed to be spread evenly over the class-interval of 5s. 0d.

$$\therefore Q_1 = \text{£5 5s. 0d.} + \left(\frac{446 - 414}{379}\right) \times \text{5s. 0d.}$$

$$= \text{£5 5s. 0d.} + \frac{32}{379} \times \text{5s. 0d.}$$

Lower Quartile wage = £5 5s. 0d. + 5·7d. = £5 5s. 6d. to nearest 1d.

1,337th item lies in group £5 15s. 0d. – £6 0s. 0d. containing 227 items;

$$Q_3 = \text{£5 15s. 0d.} + \left(\frac{1{,}337 - 1{,}267}{227}\right) \times \text{5s. 0d.}$$

$$= \text{£5 15s. 0d.} + \frac{70}{227} \times \text{5s. 0d.}$$

Upper Quartile wage = £5 15s. 0d. + 18·5d. = £5 16s. 6d. to nearest 1d.

Apart from their value in providing a description of any distribution, the quartiles and deciles are especially useful for comparison of two distributions, *i.e.*, contrasting the values in each series at the lower quartile position, upper quartile and so on. As explained above, the values of the quartiles or any of the deciles can be estimated from the ogive, as was the Median in Fig. 14. Neither the quartiles nor deciles are averages, they are measures of dispersion and as such are discussed in the next chapter. They are discussed at this stage simply because they are

[1] When the series is ungrouped, the formula for Q_1 is $\frac{N+1}{4}$, and for Q_3 $\frac{3(N+1)}{4}$. If the division yields an odd quarter in the quotient, the answer may be expressed to the nearest unit.

derived by the same methods as those employed in calculating the Median.

The Mode

Statements such as 'the average man prefers this brand of cigarettes', or that 'the average woman uses cosmetics', are frequently read and overheard. Used in this context, the term 'average' means the majority and not the Arithmetic Mean discussed earlier in this chapter. The fact that the Mean does not always provide an accurate reflection of the data due to the presence of extreme items has already been stated; similarly, the Median may prove to be quite unrepresentative of the data owing to an uneven distribution of the series. For example, the values in the lower half of a distribution range from, say, £20 to £100, while the same number of items in the upper half of the series range from £100 to £5,000 with most of them nearer the higher limit. In such a distribution the Median value of £100 will provide little indication of the true nature of the data.

Both these shortcomings are overcome by the use of the Mode, which refers to the value which occurs most frequently within a distribution. This particular 'average' is the easiest of all to find in some distributions,[1] since it is the value corresponding to the largest frequency. Thus in the following distribution which is discrete:

No. of Rooms ..	1	2	3	4	5	6	7	8	9	10	11
Frequencies ..	4	9	15	19	24	38	26	18	13	7	1

the modal value or mode is '6', since it appears more times in the series than any other value. The Mode is a particularly useful average for discrete series, *e.g.*, number of people wearing a given size shoe, or number of children per household, etc.

The Mode by Interpolation

Ascertaining the Mode is not always quite so easy, although it is seldom necessary to find it exactly. When, as is frequently the case, it has to be located in a grouped frequency distribution, the Mode lies within a given class, *i.e.*, within the limits of the

[1] Those distributions which reveal a marked tendency to cluster around a central value, clearly indicating the Mode.

maximum and minimum values of that class. The simplest course is to select the mid-point of that particular class; this is no more arbitrary than computing the Mean from a grouped frequency distribution by multiplying the frequencies by the mid-points of the corresponding classes. As was pointed out, if the distribution were evenly dispersed throughout its range of values, the result from calculating the Mean by this arbitrary method should correspond with the Mean derived from a detailed computation. Unfortunately, such distributions are infrequent, and in consequence an alternative method has been devised to locate the Mode wherever its position is at all indeterminate.

The assumption that the frequencies in a given class are spread evenly over all the values within the limits of that class is arbitrary but, as stated above, provides in many cases a fair enough approximation to the truth. In some distributions there are more items below the modal value than above it, *e.g.*, in the classes below the modal group in value, the number of frequencies may be far greater than the number of frequencies contained in the classes in the upper regions of the table. Such a case is illustrated in Table 23, below:

TABLE 23

Commission Payments for January 1958	No. of Salesmen
£10 and under £15	6
£15 „ „ £20	12
£20 „ „ £25	30
£25 „ „ £30	53
£30 „ „ £35	77
£35 „ „ £40	96
£40 „ „ £45	54
£45 „ „ £50	37
£50 „ „ £55	19
£55 „ „ £60	8

Here, it is found that more salesmen (77) were in the class below the modal class (£35 – 40) than in the one above, containing 54. Because of this, it is likely that the concentration of the salesmen within the modal class (£35 – 40) is more marked between, say, £35 to £37 10s. 0d. than between £37 10s. 0d. and £40. In other words, had a different class interval been selected for this frequency distribution, it is possible that instead of most

frequencies being within the class £35 – 40, yielding an arbitrary Mode of £37 10s. 0d., *i.e.*, the mid-point; the Mode would have been located in a new group, say, £35 – £37 10s. 0d., yielding a modal value of £36 5s. 0d. Such a breakdown of the distribution into smaller or different groups is not possible unless the full data are given elsewhere, and in passing it may be noted that the Mode can be markedly affected by the classification adopted in compiling the grouped frequency distributions.[1] The above theory underlies the following formula for estimating the Mode which involves a simple exercise in proportions:

$$\text{Mode} = L + \left[\frac{fa}{fa + fb}\right] \times C.I.$$

Where:

fa = frequencies in group following the modal group (54).
fb = frequencies in group preceding the modal group (77).
$C.I.$ = class interval (£5).
L = lower limit of modal group (£35).

Thus:

$$£35 + \left[\frac{54}{54 + 77}\right] \times £5$$

$$= £35 + \left(\frac{54}{131}\right) \times £5$$

$$= £35 + 2 \text{ to nearest } £ = £37.$$

In this example, the use of the formula does not result in the same modal value as may be obtained by simply taking the mid-point of the modal class, *i.e.*, £37 10s. 0d. This arises because the relative sizes of the two classes adjacent to the modal class, *i.e.*, 77 and 54, are unequal. The principle of the theory may be tested by substituting more closely similar figures, *e.g.* 70 and 68 in place of 77 and 54 respectively. The Mode derived by using the above formula is then £37 10s. 0d. to the nearest 5s. 0d., the same as the mid-point.

[1] If a smooth frequency curve of this distribution could be drawn from the data available, the apex of the curve, representing the Mode, would lie to the left of a perpendicular drawn from the mid-point of the modal group limits. The only really satisfactory method of ascertaining the Mode is by deriving such a curve by mathematical methods and fixing its highest point, but the simple method described in the text gives a good enough estimate for most purposes where the data forms a humped-backed frequency curve.

Selecting the 'Average'

Most readers of the last few pages will agree that the actual process of computing any one of the three 'averages' is elementary. It is not always quite so easy, however, to decide in practice on one particular average to represent a distribution rather than any other, but provided the various advantages and disadvantages of each average are understood the problem of choice is simplified. It should be mentioned in passing that if all three averages in a series are identical, no problem arises, but as will be shown later, this identity of the 'averages' is rather unusual and, as far as data relating to social and economic affairs are concerned, practically unknown.

A summary of the characteristics of each average is given below:

Arithmetic Mean

1. Every value in the series is included, but extreme items may have a disproportionate effect on the Mean and reduce its usefulness as a summary of the whole.
2. When there are no unusual or very extreme values which would distort the Mean, then the Mean has the advantage of being representative and of being based upon every value in the distribution.
3. As an exact or computed figure it is suitable for further mathematical treatment.
4. It is simple to compute and is the best understood average.
5. The Mean multiplied by the number of values in the distribution gives their aggregate.

Median

1. Unlike the Mean, the Median may be determined where the data are incomplete, *e.g.*, irregular class-intervals and open-ended final classes.
2. Provided the number of frequencies or items in an ungrouped series is uneven, the Median will actually be one of the series as it will be if a grouped distribution contains an even number of frequencies. Otherwise, the Median is a derived figure. In contrast, the Mean seldom conforms to any individual item.

3. The value of the Median is severely limited unless it can be supplemented by other values, *e.g.*, it provides the size of the middle item only, but is independent of the range of the series, or the spread of values above or below it.
4. Under the circumstances, it will be appreciated that the Median is best used when the series is continuous, or where a discrete series may be treated as continuous. Where there is a tendency for the frequencies to cluster evenly around the middle of the series, rather than dispersing themselves unevenly throughout with clusters around the maximum and minimum values, the Median is also reliable.
5. In practice, it will not be encountered as the sole 'average' used to represent the series, but is usually compared with the Mean or Mode.

The Mode

1. The Mode has the great advantage that as it is usually an 'actual' value, it indicates the precise value of an important part of the series, but not necessarily the major part. This assumes the modal value is apparent from simple observation of the distribution, *i.e.*, an obvious concentration of frequencies around a certain value. If the Mode has been interpolated by formula or the mid-point of a large class-interval used, then the foregoing statement does not hold.
2. Unless the number of frequencies is reasonably large and the distribution reveals a marked tendency to group around a given value, the Mode is not easy to determine.[1] Such grouping is more apparent where the independent variable is discrete and it is with these that the Mode is most useful.
3. It does not lend itself as does the Mean to further mathematical treatment.
4. Like the Median, however, the Mode is unaffected by the dispersion of the series, *i.e.*, its distribution over the range. Unlike the Mean, it is not affected by extreme items.

From the above summary of the strength and weaknesses of the various averages it becomes apparent that there is no such thing as an all-purpose average. Each has its own virtues.

[1] The difficulties of ascertaining the Mode are greater than may be realised from a reading of the preceding pages, hence its greater use with data from the natural sciences, where large variations from a norm are the exception rather than the rule.

Clearly then, the choice of the average in any given case must be determined by the nature of the data and the purpose to be served by the average. If it is not forgotten that a single average is designed to replace the detail, yet at the same time to provide the outline of that detail, then the selection of the average will be seen to depend on which measure fulfils this requirement most adequately. Since the three 'averages' comprise rather different concepts, the data may be such as to warrant the use of all three, and as will be shown in the next chapter, the relationship between the three measures may be significant. In any case, the chief use of averages is to compare those of one series with the same averages of another but comparable series. In practice, the Mean is a firm favourite in so far as it is so readily computed and understood; generally speaking, it should be used instead of the others. But either the Median or even the Mode will be preferable if the generalisation concerning mid-points in the calculation of the Mean is unjustified, or the Mean is seriously affected by extreme items.

Geometric Mean

The 'averages' so far discussed, *i.e.*, Mean, Mode and Median, are important for comparison of most distributions. Less significant for this purpose, but of prime importance in the preparation of index numbers[1], is the *Geometric Mean*. This is usually given its full name, to distinguish it from the 'Mean' which is the term usually applied to the Arithmetic Mean.

The arithmetic mean is calculated by *aggregating* the values in a distribution and dividing them by their number or frequencies.

The geometric mean is derived by *multiplying* together all the values and then extracting the relevant root of the product of those values. Thus, for the following series, 4, 6, 9, the geometric mean is: $\sqrt[3]{4 \times 6 \times 9} = \sqrt[3]{216} = 6$. The root to be calculated depends on the number of values in the series; thus, with three values, it is the cube root. The principle can be summarised as follows: 'the G.M. is the nth root of the product of n items.'

It will be apparent that if n is a large number, even a dozen values in the series, the problem of computing the twelfth root of the product of the twelve values by simple arithmetic is likely to

[1] Discussed in Chapter XVI.

be a tedious operation, if not impossible. Thus for calculating the G.M., logarithms are used.

By obtaining the logarithm of each value, and aggregating the logarithms, the same purpose is being served as by multiplying the original numbers together. Having aggregated the logarithms, their sum is divided by the number of items, *i.e.*, *n*. This in turn has the same effect as calculating the *n*th root of the product of the original values. The quotient of the sum of the logarithms divided by *n* is then looked up in the tables of anti-logarithms, which yield the equivalent in ordinary values, the G.M. of the original series. In other words, find the arithmetic mean of the logarithms and convert the answer back into natural numbers.

The following examples illustrate the procedure. Example 1 illustrates the calculation of the G.M. of a simple series of un-weighted values; the second shows the calculation for an un-grouped weighted series, as might arise in the calculation of a simple index number.

(1) Calculate the G.M. of the following series: 20, 58, 87, 130, 170, 250;

$$\sqrt[6]{\quad 58 \times 87 \times 130 \times 170 \times 250.}$$

Values	*Logs*
20	1·3010
58	1·7634
87	1·9395
130	2·1139
170	2·2304
250	2·3979
	11·7461

$$\text{G.M.} = \text{Anti-log of } \frac{11{\cdot}7461}{6} = \text{anti-log of } 1{\cdot}95768.$$

which converted into original units = 90·7, or 91 to nearest unit.

(2) Calculate the G.M. of the following weighted frequency distribution:

Indices	110	125	92	100	160	84
Weight ..	4	1	3	10	5	8

$$= \sqrt{110^4 \times 125^1 \times 92^3 \times 100^{10} \times 160^5 \times 84^8}$$

Indices	*Frequency or Weight*	*Logs of Indices*	*Weight × Logs*
110	4	2·0414	8·1656
125	1	2·0969	2·0969
92	3	1·9638	5·8914
100	10	2·0000	20·0000
160	5	2·2041	11·0205
84	8	1·9243	15·3944
	31		62·5688

$$\text{G.M.} = \text{Anti-log of } \frac{62{\cdot}5688}{31} = \text{anti-log of } 2{\cdot}0183.$$

G.M. = 104·3, or 104 to nearest unit.

As with the Arithmetic Mean, so with the G.M., the formula is often given in algebraic notation. Thus with the G.M. described as the Nth root of the product of N values the usual form is:

$$g = \sqrt[N]{x_1 \times x_2 \times x_3 \ldots x_N}$$

if weights are to be introduced as in the second example above.

$$g = \sqrt[w]{x_1^{w1} \times x_2^{w2} \times x_3^{w3} \ldots x_n^{wn}}$$

where w = the total of weights used and $w^{1 \cdots n}$ are individual weights.

Since the computation involves logarithms, the first form may be written:

$$\log g = \frac{\log x_1 + \log x_2 + \log x_3 \ldots + \log_n}{N}$$

$$= \frac{\Sigma \log x}{N} \text{ or } \frac{1}{N}\Sigma (\log x)$$

The weighted series is then written:

$$\log g = \frac{w_1 \log x_1 + w_2 \log x_2 + w_3 \log x_3 + \ldots w_n \log x_n}{\Sigma w}$$

$$= \frac{\Sigma w \log x}{\Sigma w}$$

where Σw represents the total weights.

The reader may care to transpose figures from the examples in the preceding pages to test the above formula. The main use of the G.M. is discussed in the Chapter on Index Numbers.

Other Distributions

It has probably been observed that in all the grouped frequency distributions used in the various examples a major part of the frequencies were concentrated around the central values

of the independent variable. If these distributions had been graphed, they would approximate in shape to the histogram illustrated in Fig. 12. The essential characteristic of such distributions is that they may be adequately described by any one of the measures of central tendency. The reader will probably have realised that where the frequency curve of this type is perfectly symmetrical, the positions on the curve of all three measures will coincide.

The importance of this type of distribution in statistical work has already been mentioned and will become more evident in the course of the next few chapters. Certainly it is the most frequent type of distribution. But there are other varieties, the first being known as the 'J' and reverse 'J'; the other as the 'U' shaped distribution. These names are derived from the shape of the curve when the distributions are graphed. Of these, the reverse 'J' is the most frequent, the other two being seldom encountered. If the data given in Table 6 (p.44) relating to the distribution of personal incomes in the United Kingdom in the fiscal year ended April 5th, 1956, were plotted, with various incomes along the base and the frequencies against the ordinate, a reverse 'J' shaped curve would be obtained. It is important to remember in connection with such distributions as these, that the mid-point method of deriving the various averages cannot be used with the same degree of confidence as for the normal humped-back distribution since the Mean is distorted by the concentration of frequencies at one end.

CHAPTER VII

MEASURES OF VARIATION

Introduction

A brief recapitulation of the content of the preceding chapters may assist the reader in understanding the purpose of the measures to be discussed in this chapter. It may be assumed that the data have been assembled in tabular form so that the initial semblance of order, so necessary to further progress, has been achieved. In the chapter on Tabulation the various forms of presenting the data in full or more usually in abbreviated form were discussed. The conclusion was drawn that, helpful as the procedure undoubtedly is in providing some indication of the nature of the data, it is still insufficient to permit rapid comparison with comparable data drawn from other but similar sources. Graphical representation, it was found, was particularly valuable in conveying rapidly, and often very effectively, an impression of the nature of the data. They greatly facilitated comparison, although in such diagrams much of the detail had to be sacrificed.

The next stage was to summarise the data from the state of tables and frequently distributions into simple figures which would indicate the outstanding features of the series. To this end four averages were discussed in the last chapter, each with its particular advantages and shortcomings. The Mean and Mode are sometimes referred to as measures of central tendency.[1] The reason will be apparent from the examples already given, since the major part of many distributions appears to concentrate around a central value with the remaining items distributed on either side of that value. It is only because of this tendency, to which further reference will be made below, that the Mode and, sometimes the Mean, have any value as representative items. If all the items in a distribution are widely dispersed and there is no tendency to concentrate around any one value, then clearly no average can adequately summarise the distribution.

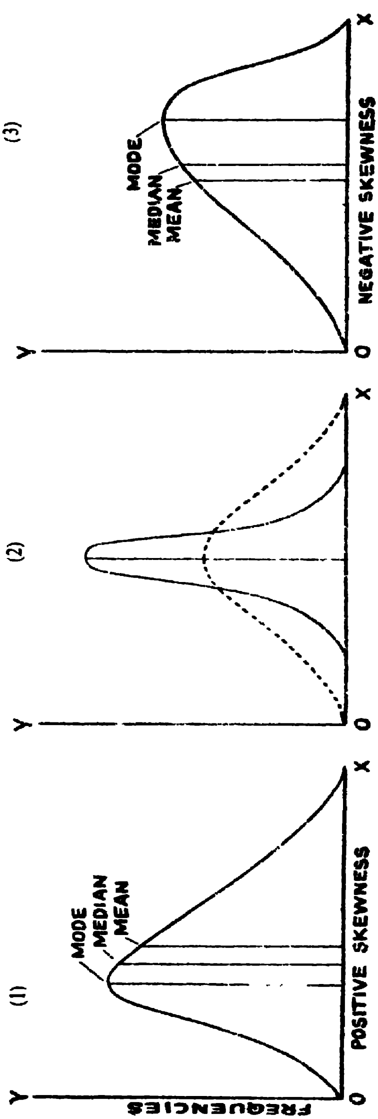
[1] For the purpose of this discussion the G.M. may be excluded.

These averages nevertheless provide only rather incomplete summaries of any frequency distribution, and important as say, the central section of any distribution may be, it is also essential to know what form the rest of the distribution takes. (Thus, if the mean age of a group of six people was 25, many varieties of combinations of ages would yield this Mean. Thus, 15, 16, 20, 22, 26 and 51 years yield a Mean of 25, as do the following: 22, 23, 24, 25, 27 and 29). The position is improved if the range, *i.e.*, the difference between the maximum and minimum values in the distribution is known. In some frequency distributions the range cannot be given, since the extreme values are unknown. Such an example is given by many income distributions in which, for example, the lower extreme is some unknown quantity 'below £500', while the upper limit is concealed in the group 'over £2,000'. When the range is given, this together with the averages, provides a good deal of information about the frequency distribution. But since the existence of a single extreme value in a distribution will greatly distort the range, its value in describing the distribution is limited. It is necessary to know how typical, *i.e.*, representative, of the distribution the average is; whether most of the values are concentrated aro that average or widely dispersed through the range. Clearly, if the intermediate values throughout the range and their distribution can be described in some numerical form, a whole series can be summarised for comparative purposes in a few simple figures. The methods used to this end produce measures of *dispersion* and *skewness*.

The Meaning of Dispersion and Skewness

To illustrate these measures, three frequency curves are shown in Fig. 15, the independent variable being plotted along the X axis and the frequencies against the ordinate on Y axis. In $_{(1)}$ the apex or peak of the curve lies to the left of centre of the 'x' axis; in $_{(3)}$, the apex is to the right of centre. The former frequency curve indicates that the majority of the frequencies are to be found around the lower values of the independent variable; in the latter, that the modal value is in the higher range of the independent variable. Fig. 15$_{(2)}$ shows two frequency curves superimposed, the continuous line is taller and narrower at the base, the dotted flatter and broader. The apex of each curve lies

Figure 15



FREQUENCY CURVES ILLUSTRATING DIFFERENT DEGREES OF DISPERSION AND SKEWNESS.

at the centre of the range of the variate, but whereas the smooth curve depicts a distribution most of whose values lie very close to the modal value (*i.e.*, given by the apex), the dotted curve depicts a distribution in which the frequencies are dispersed fairly evenly over the range of the variable. For both these curves the distribution (given by the shape of the curve) is identical on either side of the apex, so that both the distributions which they portray will have the same Mean, but the range for the 'dotted' distribution will be greater than that of the other. Such curves as these, in Fig. 15 $_{(2)}$, are described as *symmetrical*, those in Figs.$_{(1)}$ and $_{(3)}$ are *skewed*, or asymmetrical.

Further inspection of the two central curves will reveal that any distribution which when plotted forms a symmetrical frequency curve (*i.e.*, the curve is of the same shape on either side of the apex), will have all three averages, the Mean, the Median and Mode equal to each other. It follows that to the extent that a frequency curve differs from symmetry, *i.e.*, it is skewed or asymmetrical, the three averages will differ from each other. As long, however, as the distribution is reasonably symmetrical, the Mode can be ascertained approximately from the other two averages by yet another method. Experience has shown that with any hump-backed distribution, the Median lies between the Mode and the Mean, usually one-third of the difference between the two measured from the Mean. As shown in Fig. 15$_{(1)}$ and $_{(3)}$ in any distribution the Mode is at the apex of the frequency curve. If the curve is skewed, the Mean is pulled towards the longer 'tail', while the Median which divides the area under the curve into equal parts, is also pulled away from the Mode, and lies nearer the Mean. From this observation the following formula to ascertain the Mode has been evolved: Mode = Mean — 3 (Mean — Median), but the reader should bear in mind that as with so many tools, it can only be used for its particular purpose with a reasonably symmetrical distribution and *at best will generally yield only an approximate result.*[1]

The Measures of Dispersion

It is now possible to return to the measures of variation and skewness which are to serve the purpose of amplifying the generally imperfect summary of any distribution provided by

[1] The reader who has not forgotten his elementary algebra will realise that the equation will also serve to give either the Mean or Median, provided both the other averages are known.

the three averages.[1] These measures are of two main types. The first is designed to measure the variation, or more accurately, the deviation of each item in the distribution from the selected measure of central tendency, usually the Mean or Median. The second group provides a measure of the degree of asymmetry in the distribution. The first are called measures of *dispersion*, *i.e.*, they measure the extent to which the individual items in a series are dispersed or distributed over the whole range. The second are known as measures of *skewness* rather than measures of symmetry or asymmetry.

The Range

The first measure of dispersion is the *Range*. This is usually defined as the difference between the smallest and largest values of a distribution or series. The difficulty of ascertaining the range where the classes at the extremes of a frequency table are 'open', has already been mentioned. Where, as is usual, the class-limits are given, by convention the range is taken as the difference between the mid-point of the first class in order of magnitude and the mid-point of the last class. Thus, in Table 27, on page 126, the range of marks awarded is 45, *i.e.*, from 13 to 58. This is quite arbitrary, since as a result of the grouping of the data the actual value of the smallest and largest items cannot be ascertained. The weakness of the range is almost self-apparent. It requires only one extreme item at either end of the series to render it virtually valueless as a reliable indication of the data. It is possible to have two distributions with the same range, but whereas in the one the frequencies are fairly evenly distributed throughout the range of the independent variable, in the other the majority of values or observations are concentrated about a single value. In brief, dependence on the two extreme items renders the Range most unreliable as a guide to the dispersion of the values within a distribution. Its chief merit lies in its simplicity.

The Quartile Deviation

The weakness of the Range can be partially overcome if a measure of dispersion is employed which covers only a restricted range of items so that any extreme values are effectively excluded. Generally, the majority of the frequencies of a frequency

[1] As the dispersion increases, so the averages become less typical or representative.

distribution is to be found in the central part; hence, it is natural that the *Quartile Deviation* should have been evolved. This measures the dispersion of that part of any series lying between the two quartiles, *i.e.*, upper and lower quartiles.

The formula for the Quartile Deviation is written $\frac{Q_3 - Q_1}{2}$

The smaller the result given by this formula, the less is the dispersion of the middle half of the distribution about the Median. This is the average normally used with this dispersion measure. The Quartile Deviation is an absolute measure which is affected by the values of the observations in the distribution, so that the Q.D. of one distribution may be much greater than that of another, although the dispersion of frequencies is in fact smaller if, for example, the values in the latter distribution are much smaller than those in the former.

This measure depends finally, like the Range, on only two derived limits, in consequence it is sometimes described as the '*semi-inter-quartile* range'. Unfortunately, it provides no indication of the degree of dispersion or grouping of the other half of the distribution lying beyond the limits of the two quartiles. Consequently, some further measure is required which will indicate the dispersion of all the items throughout their range.

The Mean Deviation

The Mean Deviation measures the *average* or mean of the sum of all the deviations of every item in the distribution from a central value (either the Mean or Median). The Mean Deviation, therefore, provides a useful method of comparing the relative tendency of the values in comparable distributions to cluster around a central value or to disperse themselves throughout the range.

The following example illustrates the basic principles underlying the Mean Deviation.

In this example the deviations are calculated from the Median value of £4.(1) The value of the Mean Deviation for any distribution is a minimum when the deviations are measured from the Median.

(1) In contrast to the computation of the true Mean, ~~shown in~~

TABLE 24

(1) Unit Values	(2) Frequency	(3) Absolute Deviations from Median (Median = £4)	(4) Products (Col. 2 × 3)
£1	3	3	9
£2	7	2	14
£3	9	1	9
£4	11	0	0
£5	11	1	11
£6	8	2	16
£7	6	3	18
	55		77

$$\text{M.D.} \quad \frac{77}{55} = £1{\cdot}40.$$

~~Table 21~~, which involved adding to an assumed Mean the average of the net total of deviations from that Mean, the signs + and — before the deviations are ignored for the purpose of computing the Mean Deviation. The reason for this arises from the fact ~~already explained~~ that the sum of the deviations of the individual values of a series from their Mean equals zero; if then the signs were to be included, no other result than zero would be

TABLE 25

(1) Values	(2) Frequency	(3) Deviations in Class Interval Units from AM = 45	(4) Fd × CI Units
10 and under 20 ..	2	—3	—6
20 ,, ,, 30 ..	4	—2	—8
30 ,, ,, 40 ..	4	—1	—4
40 ,, ,, 50 ..	8	0	0
50 ,, ,, 60 ..	6	1	6
60 ,, ,, 70 ..	3	2	6
70 ,, ,, 80 ..	2	3	6
	29		0

$$\text{M D. ignoring signs} = \frac{36}{29} \times \text{C.I.} = 1{\cdot}24 \times 10 = 12{\cdot}4.$$

obtained. By treating all the deviations as positive, the Mean Deviation is 12·4. The foregoing example illustrates the method.

The signs may justifiably be ignored, since this measure is not designed to reveal the *manner* in which the values are distributed about their Mean. In other words, it is not designed to indicate how many of the frequencies lie below the Mean or Median, or above it. That is indicated by the skewness measures. The M.D. is concerned with the *extent* to which the values are dispersed about the Mean, regardless of whether the majority of values are greater than the Mean or smaller.

The computation of the Mean Deviation is much more complicated where the Mean or the Median proves to be an awkward number entailing tedious arithmetical calculation in deriving the products of the deviations and their respective frequencies. In such cases an arbitrary origin or assumed mean is selected and the adjustment is made at the end on lines similar to the correction made when calculating the Mean from an arbitrary origin as on p. 99. This difficulty usually arises with grouped distributions in which the Mean proves to be an awkward fraction for further calculations. Often the mid-point of the middle group is arbitrarily selected as the value,[1] from which the deviations may be computed. These points will be illustrated in the examples showing the computation of the next measure of dispersion. The rather complex method of calculating the M.D. from an arbitrary origin is not dealt with here, since this particular measure of dispersion is little used.

The Standard Deviation

This is by far the most important of the dispersion measures. The Mean Deviation is nowadays only of academic interest, and in practice has been replaced by the Standard Deviation, which enters into so many of the advanced formulae.

From the point of view of the mathematician, the practice of ignoring the signs before the deviations when computing the Mean Deviation is quite unjustifiable, and in consequence the Mean Deviation is unsuitable for use in further calculation. On the other hand, to leave the signs in, will, as has already been pointed out, reduce the Mean Deviation to zero. The Standard Deviation overcomes this problem by 'squaring' the deviations.

[1] Termed the 'arbitrary origin' or the 'assumed' mean.

TABLE 26

CALCULATION OF THE STANDARD DEVIATION

Classes	f	Deviation from A.M. (45)	$f \times d$	$(fd \times d) = fd^2$
10 and under 20 ..	2	— 30	— 60	1,800
20 " " 30 ..	4	— 20	— 80	1,600
30 " " 40 ..	4	— 10	— 40	400
40 " " 50 ..	8	0	0	0
50 " " 60 ..	6	10	60	600
60 " " 70 ..	3	20	60	1,200
70 " " 80 ..	2	30	60	1,800
	29		— 180 + 180	7,400

$$\text{S.D.} = \sqrt{\frac{\text{Sum of frequencies} \times \text{deviations squared}}{\text{Sum of frequencies}}} = \sqrt{\frac{\Sigma f d^2}{\Sigma f}}$$

$$\text{S.D.} = \sqrt{\frac{7{,}400}{29}}$$

S.D. $= \sqrt{255{\cdot}172} = 16$ (answer rounded to nearest unit since the original data do not justify greater accuracy).

Thus, $(-2)^2$ is 4, just as the square of $+2$ is 4. As with the Mean Deviation, the sum of the products is divided by the total frequencies. The mean of the sum of the squared deviations is known as the *variance*, but before it can be related to any other statistic, *e.g.*, the mean, the square root of the variance must be obtained. This is known as the Standard Deviation. The example above indicates the principle; the reader should compare the result with the Mean Deviation computed from the same data above in Table 25. He will note that the Standard Deviation is larger, owing to the fact that the process of squaring gives relatively greater emphasis to the extreme values in the distribution.

Short Method of calculating the Standard Deviation

The deviations are usually measured from the Mean of the distribution, but if, as is often the case, the Mean is not a round number coinciding with say the value of the mid-point of any class, then the calculations could be extremely tedious. Consequently, a method has been evolved to avoid this, which in its essence is the same as the short method for calculating the Mean

itself.[1] It will be remembered that an assumed Mean was chosen and from it the differences for each class worked out in terms of class intervals and the residual or net difference converted and deducted from or added to the value of the assumed mean. To

TABLE 27

EXAMINATION MARKS AWARDED TO 392 CANDIDATES

Classes	Mid-points (2)	Frequencies (3)	Deviations (4)	$f \times d' = fd'$ (5)	$fd' \times d' = fd'^2$ (6)
11 .. 15	13	6	— 5	— 30	150
16 .. 20	18	12	— 4	— 48	192
21 .. 25	23	30	— 3	— 90	270
26 .. 30	28	53	— 2	— 106	212
31 .. 35	33	77	— 1	— 77	77
36 .. 40	38	96	0	0	0
41 .. 45	43	54	1	54	54
46 .. 50	48	37	2	74	148
51 .. 55	53	19	3	57	171
56 .. 60	58	8	4	32	128
		392		— 134	1,402

calculate the Standard Deviation the same procedure is employed with the additional step of multiplying the *products of* the frequencies and deviations, by the deviations. This is the 'squaring' of the deviations required for the Standard Deviation. Then, as for the A.M., a correction is introduced to derive the Standard Deviation of the distribution from the true mean. Table 27 provides a simple example, the successive steps in the calculations are detailed below.

1. The selection of the mid-points of the classes is relatively simple. The series is discrete, since the individual values are determinate amounts, *i.e.*, one unit is the minimum variation. The mid-points are derived by halving the sum of the limits of the individual class.

2. The deviations are measured from an assumed origin (38, the mid-point of 36-40 group), instead of the true Mean, which is unknown, and a correction will have to be introduced later to offset the discrepancy this method will introduce into the calculation.

[1] If the student reader has forgotten the process he should refresh his memory by reference to pp. 97-99.

3. The deviations from the assumed origin are expressed, as can be seen in Col. 4, in terms of class intervals (multiples of 5) and the result will be expressed in these terms.
4. Col. 5 gives the products of the frequencies and the deviations (Col. 3 × Col. 4). This column is important since from it is derived the correction to adjust the error introduced by calculating the Standard Deviation from an arbitrary origin. The net products will also provide the fraction for computing the Mean, thus the *true* Mean equals $38 + \left(\frac{-134}{392}\right) \times 5 =$ $38 - 1{\cdot}7$, or 36 to the nearest unit.
5. Col. 6 gives the results of multiplying Col. 5 (*i.e.*, products of frequencies and deviations) by the deviations. This can be verified by squaring all the deviations in Col. 4 and multiplying the squared products by their respective frequencies in Col. 3. The results will be the same as given by the method shown in Col. 6.

The Standard Deviation from the assumed mean can now be calculated by dividing the total of Col. 6 by the total frequencies and extracting the square root of the quotient, and from this the correction for the use of an assumed Mean is taken. The student should note that the correction fraction $\left(\frac{-134}{392}\right)$ is the same for the A.M. as it is for the S.D. except that for the latter calculation it is squared. The calculations are as follows: the usual symbol for the Standard Deviation is σ.[1]

$$\sigma^2 = \left(\frac{1{,}402}{392}\right) - \left(\frac{-134}{392}\right)^2$$

$$\sigma = \sqrt{\left(\frac{1{,}402}{392}\right) - \left(\frac{-134}{392}\right)^2}$$

It should be noted in passing, that the correction fraction should be squared and subtracted from the other fraction, before the square root is calculated.

$$\sigma = \sqrt{3{\cdot}5765 - ({\cdot}3418)^2} = \sqrt{3{\cdot}5765 - {\cdot}1169}$$

$$= \sqrt{3{\cdot}4596} = 1{\cdot}86$$

But σ is still in class-interval units, to convert to the original units 1·86 is multiplied by 5; then σ in original units $= 9{\cdot}3$ or 9 to nearest whole unit.

[1] The usual symbol for the Standard Deviation is σ (little sigma), but S.D. will serve equally well.

The Standard Deviation is the square root of the average of the squared deviations measured from the Mean. It is sometimes described as the root-mean-square deviation, although the only advantage to be derived from this name is that it indicates the method of calculation.

Characteristics of Dispersion Measures

The major characteristics of the measures of dispersion may be summarised as follows:

Range

1. The simplest to derive and the easiest to comprehend.
2. Its value as an indication of the variation in the data may be virtually nullified by the existence of one exceptionally large or small value.
3. It provides no indication as to the distribution of the frequencies between the limits of the range.

Quartile Deviation

1. This measure is not difficult to calculate, but covers only half the items within the distribution. It does eliminate, however, the risk of extreme items which may seriously distort the Range.
2. As with the Range, its value is based on the values of the two limits, *i.e.*, Q_1 and Q_3 with all the attendant disadvantages arising from this fact.
3. It bears no relationship to any fixed point in the distribution as do the M.D. and S.D.; nor is it affected by the distribution of the individual values lying between the quartiles.

Mean Deviation

1. Unlike the Q.D. it is affected by every value in the distribution.
2. It indicates the extent of the deviation of all values from a given value, in this case the Median or Mean of the distribution.
3. For the purpose of further mathematical treatment the M.D. is unsatisfactory.

Standard Deviation

1. Like the M.D., it includes every value of the distribution.

2. It is itself the result of correct mathematical processes and thus further calculations may be based upon it.
3. It is the best measure of dispersion and, as will be seen in the later chapters, is of very great importance for sampling theory.

Coefficient of Variation

It cannot be sufficiently emphasised that all measures of dispersion are in terms of the units in which the original values are expressed. Thus, the S.D. of men's heights, weight of cotton bales and salesmen's salaries will be expressed in inches, hundredweights and pounds sterling, respectively. These measures of absolute variation cannot be compared with each other, if expressed in differing units, or if the average values of two distributions in a comparable field are widely dissimilar. Thus, in the case of differing units, if the A.M. and S.D. of one distribution are expressed in centimetres, and for the other in feet, the units must either be converted to a common base, *e.g.*, both series in feet; or a standard measure devised which ignores the original units of measurement. Similarly, where the means of the distributions are widely dissimilar, *e.g.*, the average levels of remuneration received by the administrative and labouring sections of a large organisation, respectively, then the dispersions within the two groups can only be compared by relating them to some 'equalising' factor.

This is done by turning the absolute measure of dispersion, *i.e.*, the S.D., into a relative measure. More precisely, the S.D. is related to some other measure directly connected with the same distribution, *e.g.*, it is frequently expressed as a *percentage* of the Mean of that distribution. This new measure is termed the *Coefficient of Variation*, normally written as $CV = \frac{100\sigma}{AM}$ Where the series being compared are expressed in the same unit of measurement and the Means are similar, no advantage is to be gained by calculating the coefficient of variation. The S.D. is then quite sufficient.

Measures of Skewness

So far, only the first group of measures of variation have been covered, those which indicate the dispersion of the frequencies throughout the range of the independent variable.

The second group was described as measuring the degree of symmetry of any distribution plotted as a frequency curve. Any curve which is not 'symmetrical' may be described as 'asymmetrical,' or '*skewed.*' The latter term is generally employed and the statistician refers to measures of skewness. Most 'hump-backed' or uni-modal frequency distributions are skewed (*i.e.*, not symmetrical); and to that extent the characteristic of a symmetrical distribution, *i.e.*, identity of values of A.M., Median and Mode, is absent.

It is a logical step, therefore, to develop some measure which shows the degree to which these three measures of central tendency diverge. The difference between them provides the first measure of skewness. This difference, however, could be unsatisfactory on two counts.

1. It would be expressed in the unit of value of the distribution and could therefore not be compared with another comparable series expressed in different units.
2. Distributions vary greatly and the difference between, say, the Mean and Mode in absolute terms might be considerable in one series and small in another, although the frequency curves of the two distributions were similarly skewed.

If the absolute differences were expressed in relation to some measure of the spread of the values in their respective distributions, the measures would then be *relative* and not *absolute* and therefore directly comparable.

The above considerations form the basis of Professor Karl Pearson's formula for deriving what is known as a coefficient of skewness:

$$sk = \frac{\text{Mean} - \text{Mode}}{\text{SD}}$$

The Mode was criticised in an earlier chapter as being particularly difficult to determine precisely for many frequency distributions. Consequently a variation of this formula is used:

$$sk = \frac{3\,(\text{Mean} - \text{Median})}{\text{SD}}$$

An alternative measure of skewness has been proposed by the late Professor Bowley. This is based on the relative positions of the Median and the Quartiles. If the distribution were symmetrical then Q_1 and Q_3 would be at equal distances from the

Median. Then it follows that $(Q_3 - Me) - (Me - Q_1) = O$. The more skewed the distribution, the larger will be the difference between these two quantities, which can be re-arranged as follows: $(Q_3 - Me) - (Me - Q_1) = Q_3 + Q_1 - 2Me$. This measure of skewness is expressed in absolute terms so that if there were two distributions, one highly skewed and the other much less, comprising very different-sized variables, *e.g.* pounds and ounces, then despite the fact that the distribution with the large values was less markedly skewed than the other, the above measure of skewness would yield a larger result. To overcome this weakness, the absolute measure is converted into a relative measure (as in the Pearson coefficient above) by relating it to the Quartile Deviation, which is itself a reflection of the absolute variation of the independent variable. Thus, we get:

$$sk = \frac{Q_3 + Q_1 - 2Me}{\frac{Q_3 - Q_1}{2}} = \frac{2(Q_3 + Q_1 - 2Me)}{Q_3 - Q_1}$$

Beyond the fact that symmetry (*i.e.*, complete absence of skewness) is indicated by 0 (zero) in both the above formulae, *i.e.* Bowley's and Pearson's, the coefficient of skewness derived by the two measures are *not* comparable.

Formulae for Measures of Dispersion

The mathematical notation of the measures of dispersion like those of the Arithmetic Mean, given in the preceding chapter, is fundamentally simple.

Only two of the four measures of dispersion are so expressed, the M.D. and the S.D.

The Mean Deviation calculated from the true Mean of an ungrouped series:

$$\text{M.D.} = \frac{\Sigma|d|}{N} \textit{ i.e., } \frac{\text{Sum of the deviations (ignoring signs)}}{\text{number of items}}$$

When the data are in the form of a frequency distribution,

$$= \frac{\Sigma f|d|}{\Sigma f} \textit{ i.e., } \frac{\text{Sum of frequencies} \times \text{deviations (ignoring signs)}}{\text{Sum of frequencies}}$$

The *Standard Deviation* is usually represented by the sign σ (small Greek letter 'sigma'), but S.D. is often used.

The S.D. from an ungrouped distribution: $\sigma = \sqrt{\dfrac{\Sigma d^2}{N}}$

The S.D. of a grouped frequency distribution computed from the true Mean:

$$\sigma = \sqrt{\frac{\Sigma f d^2}{N}} \quad \text{Equally one can write: } \sigma = \sqrt{\frac{\Sigma f\, d^2}{\Sigma f}}$$

since Σf and N both refer to the total frequency.

If the S.D. is computed from an assumed Mean or arbitrary origin, then:

$$\sigma = \sqrt{\frac{\Sigma f\,(d')^2}{N} - \left(\frac{\Sigma f\, d'}{N}\right)^2}$$

Where the symbol d' denotes deviations from arbitrary origin or assumed mean.

If the calculation has been performed with the deviations expressed in group intervals:

$$\sigma = \sqrt{\frac{\Sigma f\, d'^2}{N} - \left(\frac{\Sigma f\, d'}{N}\right)^2} \times i$$

where i = group interval.

The same comment applies here as was made in connection with the formulae for the various averages. If these processes are really understood, these formulae can always be constructed. Nevertheless, most students prefer, sometimes unwisely, to rely on their memories.

Conclusions

The student has now reached the end of what may be termed 'descriptive statistics'. After collection of the data, all the processes described so far have been in the nature of summarizing with the object of making simple comparisons. Such comparisons can be made as between two or more distributions by means of simple tables, or by use of diagrams as was shown in Chapter V. The next stage was to summarise the data still further by means of 'averages' and at the beginning of this chapter (p. 117) it was explained that such averages may describe a distribution very imperfectly. In fact, two widely disparate distributions can have identical means. but whereas in the one case the range of the values is very slight, in the other it is considerable. Hence, it is essential to calculate statistics which will

measure not merely averages, but measure and describe the degree of dispersion. In statistics, it is not the tendency for many observations or values to conform to an average that is significant; rather it is the tendency for values to deviate from the norm which interests the statistician even more. The importance of variations about an average will be discussed in the next section of this book.

A Worked Example

It may help the reader if at this stage a complete worked example giving the various averages and measures discussed so far can be traced through the various stages.

The figures in the table below relate to the employed male population classified according to their age.

TABLE 28

ANALYSIS BY AGE OF EMPLOYED MALES IN GREAT BRITAIN, MAY 1957 (THOUSANDS)

AGE	Great Britain	London and S.E. England
19 years and under	1,094	211
20 " " " 24 years	1,237	286
25 " " " 29 "	1,512	359
30 " " " 34 "	1,594	382
35 " " " 39 "	1,546	369
40 " " " 44 "	1,515	373
45 " " " 49 "	1,560	398
50 " " " 54 "	1,473	374
55 " " " 59 "	1,192	291
60 " " " 64 "	880	212
65 years and over	597	164
TOTAL	14,200	3,419

Source: Ministry of Labour Gazette, April, 1958.

The calculations are set out in full below for 'Great Britain' and 'London and S.E. England' with a step by step commentary on the first set of figures. In particular the selection of mid-points should be noted. The first class is assumed to have a lower limit of 15 years, since this is the school leaving age. When, as is often the case, the lower limit is not given for the first class, the statistician must make the best estimate he can. For example, data from the 1953/4 Household Expenditure Enquiry were classified with the first group 'Under £3'. It would be illogical to assume that the

mid-point was 1½, based on an interval of £0-3. Even the poorest household probably has at least £2 each week so that the best estimate of the mid-part for that class would be £2½.

Returning to the data in the above table, it will be seen that each class is written as, *e.g.*, 20 and under 24 years. The upper limit of this class is for all practical purposes 25 years less one day, *i.e.*, a male whose birthday falls on the day after this registration is classified as 'under 25'. We treat age in this case as a continuous variable so the mid-points will be

$$\frac{20 + 25}{2} = 22 \cdot 5$$

i.e., the sum of the lower limits of two successive classes. Had the age of each male been given to the nearest year, *e.g.*, age on last birthday, then the classification used would have been 20–24, 25–29, and the distribution would be discrete. In this case, the mid-points would be 22, and 27 years. This can be demonstrated quite simply by setting down all the possible ages classified between 20–24, , *i.e.*, 20, 21, 22, 23 and 24; of which 22 is the mid-value. Since the above variable, however, is continuous and the ages classified in the class '20 and under 25' range from 20 years exactly to 24 years 364 days, the mid-point is 22½ years. In the first class the mid-point is $\frac{15 + 20}{2} = 17\frac{1}{2}$ years.

TABLE 29

DATA FOR GREAT BRITAIN FROM TABLE 28

Age	No of employed males	Mid-point	d' ÷ CI	fd' ÷ CI	f(d')² ÷ CI	Cum.f.
19 years and under	1,094	17½	—5	— 5,470	27,350	1,094
20 " " " 24 years	1,237	22½	—4	— 4,948	19,792	2,331
25 " " " 29 "	1,512	27½	—3	— 4,536	13,608	3,843
30 " " " 34 "	1,594	32½	—2	— 3,188	6,376	5,437
35 " " " 39 "	1,546	37½	—1	— 1,546	1,546	6,983
40 " " " 44 "	1,515	42½	0	— 19,688	0	8,498
45 " " " 49 "	1,560	47½	+1	+ 1,560	1,560	10,058
50 " " " 54 "	1,473	52½	+2	+ 2,946	5,892	11,531
55 " " " 59 "	1,192	57½	+3	+ 3,576	10,728	12,723
60 " " " 64 "	880	62½	+4	+ 3,520	14,080	13,603
65 years and over	597	67½	+5	+ 2,985	14,925	14,200
TOTAL	14,200			(+ 14,587 — 19,688) — 5,101	115,857	

For all the other classes but the last in Table 29 the class interval is five years. The final class, however, is open-ended but it is treated as having the same class interval on the assumption that there are probably roughly the same numbers of men over $67\frac{1}{2}$ years employed as between 65 and $67\frac{1}{2}$. The choice may either over- or under-estimate the numbers of males over $67\frac{1}{2}$ years still employed, but as the number in this class is *relatively* small, any error in the assumption is likely to have a very slight effect on the final answer.

Working on the figures for Great Britain, the first calculation provides the Arithmetic Mean. It is simplest to use the short-cut method of working in deviations from the assumed mean ($42\frac{1}{2}$ years, *i.e.*, the mid-point of the class 40 and under 44 years) measured in group intervals:

$$
\begin{aligned}
AM &= 42\tfrac{1}{2} + \left(\frac{-5{,}101}{14{,}200}\right) \times 5 \text{ years} \\
&= 42\tfrac{1}{2} + 5\,(-0{\cdot}36) \text{ years} \\
&= 42\tfrac{1}{2} - 1{\cdot}8 \text{ years} \\
&= 41 \text{ years (to nearest year)}
\end{aligned}
$$

The Standard Deviation is derived from the formula:

$$\sqrt{\frac{\Sigma f d'^2}{N} - \left(\frac{\Sigma f d'}{N}\right)^2} \times i$$

i.e., working in class intervals from the assumed Mean.

Substituting the true values for the symbols in the above formula

$$
\begin{aligned}
S.D. &= \sqrt{\frac{115{,}857}{14{,}200} - \left(\frac{-5{,}101}{14{,}200}\right)^2} \times 5 \\
&= \sqrt{8{\cdot}1589 - (0{\cdot}36)^2} \times 5 \\
&= \sqrt{8{\cdot}1589 - 0{\cdot}1296} \times 5 \\
&= \sqrt{8{\cdot}0293} \times 5 \\
&= 5\,(2{\cdot}834) = 14{\cdot}17 = 14 \text{ years to nearest year.}
\end{aligned}
$$

The Median from a grouped frequency distribution is derived by the formula $\frac{N}{2}$. Thus:

$$\frac{14{,}200}{2} = 7{,}100\text{th item which lies in the group 40–44 years.}$$

Thus the value of the Median by interpolation

$$= 40 + \left(\frac{7,100 - 6,983}{1515}\right) \times 5 \text{ years}$$

$$= 40 + \left(\frac{117}{1,515}\right) \times 5$$

$$= 40 + 5\,(0{\cdot}077) = 40 \text{ years } 5 \text{ months.}$$

$$= 40 \text{ years to nearest year.}$$

The values at the Quartiles are derived in the same way

$$Q_1 = \frac{14,200}{4} = 3,550\text{th item}$$

$$= 25 + \left(\frac{3,550 - 2,331}{1,512}\right) \times 5$$

$$= 25 + \left(\frac{1,219}{1,515}\right) \times 5$$

$$= 25 + 5\,(0{\cdot}805) = 29 \text{ years to nearest year.}$$

$$Q_3 = \frac{3(14,200)}{4} = 10,650\text{th item}$$

$$= 50 + 5\left(\frac{10,650 - 10,058}{1,473}\right)$$

$$= 50 + 5\left(\frac{592}{1,473}\right) = 50 + 5\,(0{\cdot}402) = 52 \text{ years to nearest year.}$$

The Quartile Deviation is easily obtained by using the above results in the formula $QD = \dfrac{Q_3 - Q_1}{2}$

$$= \frac{52 - 29}{2} \text{ years} = \frac{23}{2} = 11\tfrac{1}{2} \text{ years}$$

The Coefficient of Variation is derived from $\dfrac{100 \times \text{S.D.}}{AM}$

$$\frac{100 \times 14{\cdot}2}{40{\cdot}75} = \frac{1,420}{40{\cdot}75} = 35\%$$

TABLE 30

LONDON AND SOUTH EAST ENGLAND

Age	No. of employed males	Mid-point	d÷CI	fd ÷ CI	f(d)² ÷ CI	Cum.f.
19 years and under	211	17½	—5	— 1,055	5,275	211
20 ,, ,, ,, 24 years	286	22½	—4	— 1,144	4,576	497
25 ,, ,, ,, 29 ,,	359	27½	—3	— 1,077	3,231	856
30 ,, ,, ,, 34 ,,	382	32½	—2	— 764	1,528	1,238
35 ,, ,, ,, 39 ,,	369	37½	—1	— 369	369	1,607
40 ,, ,, ,, 44 ,,	373	42½	—0	— 4,409	0	1,980
45 ,, ,, ,, 49 ,,	398	47½	+1	+ 398	398	2,378
50 ,, ,, ,, 54 ,,	374	52½	+2	+ 748	1,496	2,752
55 ,, ,, ,, 59 ,,	291	57½	+3	+ 873	2,619	3,043
60 ,, ,, ,, 64 ,,	212	62½	+4	+ 848	3,392	3,255
65 years and over ..	164	67½	+5	+ 820	4,100	3,419
TOTAL	3,419			(+ 3,687 — 4,409) — 722	26,984	

The calculations follow the same pattern as for the first distribution and no comment is required.

$$AM = 42\tfrac{1}{2} + \left(\frac{-722}{3{,}419}\right) \times 5 = 42\tfrac{1}{2} + (-0{\cdot}21)\,5$$

$$42\tfrac{1}{2} - 1{\cdot}05 = 41 \text{ years to nearest year.}$$

$$SD = \sqrt{\frac{\Sigma f(d')^2}{N} - \left(\frac{\Sigma f d'}{N}\right)^2} \times i$$

$$= \sqrt{\frac{26{,}984}{3{,}419} - \left(\frac{-722}{3{,}419}\right)^2} \times 5$$

$$= \sqrt{7{\cdot}8924 - (-0{\cdot}21)^2} \times 5$$

$$= \sqrt{7{\cdot}8924 - 0{\cdot}0441} \times 5$$

$$= \sqrt{7{\cdot}8483} \times 5$$

$$= 5\,(2{\cdot}801) = 14{\cdot}005 = 14 \text{ years to nearest year.}$$

$$\text{Median} = 1{,}709{\cdot}5 = 40 + 5\,\frac{1{,}709{\cdot}5 - 1{,}607}{373} = 40 + 5\,\frac{102{\cdot}5}{373}$$

$$= 40 + 5\,(0{\cdot}275) = 41 \text{ years.}$$

Quartile Deviation:

$$\text{Position of } Q_3 = \frac{3(3{,}419)}{4} = \frac{10{,}257}{4} = 2{,}564\tfrac{1}{4} \text{ rank}$$

Value of $Q_3 = 50 + \frac{5(2{,}564\frac{1}{4} - 2{,}378)}{374}$

$= 50 + \frac{5(186\frac{1}{4})}{374} = 50 + \frac{931\frac{1}{4}}{374} = 50 + 2{\cdot}489$

$=$ 52 years to nearest year.

Position of $Q_1 = \frac{3{,}419}{4} = 854\frac{3}{4}$ rank

Value of $Q_1 = 25 + \frac{5(854\frac{3}{4} - 497)}{359}$

$= 25 + \frac{5(357\frac{3}{4})}{359} = 25 + \frac{1{,}788\frac{3}{4}}{359} = 25 + 4{\cdot}98$ years

$=$ 30 years

Quartile Deviation $= \frac{Q_3 - Q_1}{2} = \frac{52\frac{1}{2} - 30}{2} = \frac{22\frac{1}{2}}{2}$

$= 11\frac{1}{4}$ years

Coefficient of Variation

$$\frac{\text{S.D.} \times 100}{AM} = \frac{14{\cdot}005 \times 100}{41{\cdot}5} = \frac{1405{\cdot}0}{41{\cdot}5} = 33{\cdot}74\%$$

It will be noted that several places of decimals have been used in the calculations above. However, in the comparison of results below all the statistics have been rounded to the nearest year as this approximation is adequate for the data involved and any greater degree of accuracy is unobtainable from data classified as in the original tables.

	Great Britain	London and S.E. England
Number in thousands	14,200	3,419
Arithmetic Mean	41 years	41 years
Median value	40 ,,	41 ,,
Lower Quartile Q_1 value	29 ,,	30 ,,
Upper Quartile Q_3 value	52 ,,	52 ,,
Quartile Deviation	12 ,,	11 ,,
Standard Deviation	14 ,,	14 ,,
Coefficient of Variation	35%	34%

CHAPTER VIII

ACCURACY AND APPROXIMATION

Most people tend to think of values and quantities expressed in numerical terms as being exact figures; much the same as the figures which appear in the trading account of a company. It therefore comes as a considerable surprise to many to learn that few statistics are exact. Many published figures are only approximations to the real value, while others are estimates of aggregates which are far too large to be measured with precision. For example, the Monthly Digest of Statistics contains many series of economic statistics which are expressed to the nearest million pounds, or hundreds of thousands of yards, or thousands of tons. It would be very satisfactory to know that every figure in that Digest was correct to the nearest unit, but in many cases it would be quite impossible to achieve complete accuracy, for some units are always missed out in a count involving hundreds and thousands of units. To achieve such accuracy would also take a great deal of time so that when finally these statistics were published they would be so much out of date as to be useless. With economic data, early publication is usually more important than precision to the last unit. If action is required on the evidence of these data, the sooner it is taken, the better. In many series, it is not so much the aggregates themselves which are of interest as the pattern of change which emerges over a period of months. A good example is provided by the quarterly figures of prospective capital expenditure.[1] Provided gross errors in collection of these data are avoided, approximate figures will serve this particular purpose of indicating trends and changes very well.

Approximate Data

The simplest way of indicating that figures are not given precisely to the last unit is to express them to the nearest 100 or 1,000; or in some cases to the nearest 100,000 or million. Take,

[1] This series is discussed in Chapter XVII.

for example, the annual mid-year estimates of the population of England and Wales. These are based upon the figure for the last census plus the net increase (decrease) in respect of births and deaths together with net migration. This pre-supposes that the census figure was accurate, and the fact that we are informed that 43,747,888 persons were enumerated in England and Wales on the night of 8th April, 1951, suggests a quite remarkable degree of accuracy. It is, however, certain that some persons were missed out in the census count. The birth and death registration system in England and Wales is very reliable but the figures for migration are not so good. Hence, the estimate of the population of England and Wales at any date after 8th April, 1951, until the next census is inevitably subject to a margin of error. For this reason, the mid-year estimates are given to the nearest 1,000 while the population of most large towns is given to the nearest 100, the two degrees of approximation reflecting the relative reliability of the published estimates.

This desire for precision is reflected in many reports on economic trends which quote figures in great detail, rather than emphasising the trends and movements reflected in the figures. For example, if the exports of a particular product rose last year to £20,879,169 from £13,998,372 in the previous year, it is much clearer to state that the value of exports rose from about £14 million to approximately £21 million; alternatively, that this year's figure of almost £21 million was half as large again as that for last year. It is important to distinguish the purposes to which such published statistics are to be put; if they are merely inserted to indicate the course of events or approximate magnitude of the variables, then rounded figures which are immediately comprehensible are infinitely preferable to the exact figures. On the other hand, if a detailed analysis is to be carried out then the results may have to be given to the nearest unit.

The practice, described above, of expressing large figures more simply, *i.e.*, by dropping the last few digits, is described as *rounding*. This is done with the mid-year estimate of the population, the total being expressed to the nearest thousand, and such rounding implies that the last digit in the rounded figure is correct. The following are other methods available. Assuming the original figure to lie within the limits of 82,500 to 83,500, we may write:

(1) 83,000 $\pm$ 500.
(2) 83,000 correct to ·6 per cent.
(3) 83,000 correct to nearest 1,000.

In order to simplify statistical tables, the practice of rounding large figures and totals is generally resorted to. Where the constituent figures in a table together with their aggregate have been so treated, a discrepancy between the rounded total and the true sum of the rounded constituent figures frequently arises. Under no circumstances should the total be adjusted to what appears to be the right answer. A note to the table to the effect that the figures have been rounded, *e.g.* to the nearest 1,000, is all that is necessary. The same remark applies to percentage equivalents of the constituent parts of a total; if they do not add to exactly 100 per cent, leave them. This has been done in the 1955 column of Table 10 (page 49). The error arises here because each regional percentage has been calculated to two places of decimals and then rounded to the nearest first place. Similarly, in Table 9 the column headed 'Net Output' does not add up exactly to the total shown, due to the fact that the values for each class have been rounded to the nearest £000.

Biased and Unbiased Errors

The rounding of individual values comprising an aggregate can give rise to what are known as *unbiased* or *biased* errors. Table 31 below illustrates this. The *biased* error arises because all the individual figures are reduced to the lower 1,000, as in column 3; or as in column 4, where they have been raised to the higher 1,000.

The *unbiased* error is so described since by rounding each item to the nearest 1,000 some of the approximations are greater and some smaller than the original figures. Given a large number of such approximations, the final total may therefore correspond very closely to the true or original total, since the approximations tend to offset each other. This is true of Column 2, which totals 75,000, the same as the actual total expressed to the nearest 1,000. It is possible even if rather unlikely, that the 'unbiased' total may be very different from the true total if most of the approximations are in one direction only, *e.g.*, a group of figures each rounded to the nearest '000' where most of them lie just above '500'. The larger the number of values, however, the less likely is the total to differ from their true aggregate since

the unbiased rounding of each figure will tend to balance out.

With the *biased* approximations, the errors are cumulative and their aggregate increases with the number of items in the series. 'Biased' errors may arise in a variety of ways. If a number of women are asked to state their ages, and an average is computed, the latter is quite likely to be lower than the 'true' average, since many of the informants will tend to understate their ages by a year or so. Likewise, an administrative officer in computing his probable staff requirements in a group of offices may, to be on the safe side, over-estimate his needs for each office. Similarly, data based on readings from an inaccurate slide-rule, thermometer, or similar measuring instrument will be consistently biased in the same direction, *i.e.*, either above or below the true figures.

Absolute and Relative Errors

Errors may be measured in two ways: *Absolutely* and *Relatively*. The *absolute* error is the arithmetic difference between the approximated figure and the original quantity. Thus, in Table 31, the absolute error in the total arising from the 'unbiased' rounding in Column 2 is – 182

The *relative* error is generally derived by expressing the absolute error as a fraction of the estimated total, *i.e.*, 75,000 Thus $\frac{-182}{75,000} = -0024$. The same calculations have been performed

TABLE 31

EXAMPLES OF ROUNDING AND BIASED ERRORS

Actual Figures (1)	Unbiased (000) (2)	Biased (Lower 000) (3)	Biased (Higher 000) (4)
17,118	17	17	18
613	1	0	1
1,253	1	1	2
8,362	8	8	9
15,443	15	15	16
7,645	8	7	8
11,759	12	11	12
10,509	11	10	11
2,480	2	2	3
TOTAL 75,182	75	71	80

for both the biased errors, where the estimated figures are adjusted downward to the nearest 1,000, *i.e.*, in Column 3 the relative error is —·0590, when the figures are adjusted upwards, the error is +·0602 (Col. 4).

Actual Absolute Error	—182	—4,182	+4,818
Actual Relative Error	$\left(\frac{-182}{75{,}000}\right) = -0{\cdot}24\%$	$\left(\frac{-4{,}182}{71{,}000}\right) = -5{\cdot}90\%$	$\left(\frac{+4{,}818}{80{,}000}\right) = +6{\cdot}02\%$
AVERAGE, $8{,}353{\cdot}\dot{5}$	$8{,}333{\cdot}\dot{3}$	$7{,}888{\cdot}\dot{8}$	$8{,}888{\cdot}\dot{8}$

Absolute Error in Average

$$\left(\frac{-182}{9}\right) = -20{\cdot}\dot{2} \qquad \left(\frac{-4{,}181}{9}\right) = -464{\cdot}6 \qquad \left(\frac{4{,}818}{9}\right) = +535{\cdot}\dot{3}$$

Relative Error in Average

$$\left(\frac{-182}{9} \div \frac{75{,}000}{9}\right) = -0{\cdot}24\% \qquad \left(\frac{-4{,}181}{9} \div \frac{71{,}000}{9}\right) = -5{\cdot}90\% \qquad \left(\frac{4{,}818}{9} \div \frac{80{,}000}{9}\right) = +6{\cdot}02\%$$

The advantage of relatives is that widely differing quantities can be compared in similar terms. Thus, an error of 100,000 in £10 million is the same as 5 in £500, *i.e.*, 1 per cent.

The *absolute* error in the aggregate of a *biased* series will tend to increase with the number of items. The *relative* error in the total figure will, generally speaking, tend to diminish as the number of items increases. If, however, every item added is biased by the same proportion as is the total before that item is added, then the relative error in the total will be unchanged.

The remainder of the calculations below Table 31 are simple, involving the calculation of the absolute and the relative errors in the *average*. The results in this case provide an indication of the accuracy, not of the aggregates themselves, but of the *averages* computed from those aggregates. The reader will note by reference to Table 31 that the relative error in both the average and the aggregate is the same. Generally, in any series where the individual figures have all been rounded to the same unit, (*e.g.*, rounding to nearest '000' means an error ± 500), then the *average* of the total is more reliable than any of the individual figures.

Errors in Calculations

Wherever any arithmetical calculation involving approximated figures is carried out, and the degree of error in those figures is known, it is possible to estimate the error arising in the final result. Starting with simple *addition:*

Add 56,000, 7,000 and 20,000 correct to 5 per cent., ·5 per cent., and ·05 per cent. respectively:

56,000	5% of which is	2,800
7,000	·5% of which is	35
20,000	·05% of which is	10
83,000		2,845

The aggregate is 83,000 $\pm$ 2,845, *i.e.*, correct to 3·43 per cent. Thus, the error in the aggregate is the sum of the absolute errors in the component items.

For *subtraction* the difference will be at a minimum if the estimate of the larger figure is assumed to be below the true figure by the amount of error, and the smaller estimated figure to be subtracted is above its true amount by the amount of error shown. A maximum result is obtained if the reverse of the above case applies. Assuming that 45,000 is to be subtracted from 72,000, and the former figure lies between 44,000 and 46,000, and the latter sum between 71,500 and 72,500, then:

Maximum difference	Minimum difference
72,500	71,500
44,000	46,000
28,500	25,500

The answer is $\frac{(28{,}500 + 25{,}500)}{2} = 27{,}000 \pm 1{,}500$, *i.e.*, correct to 5·55 per cent. With subtraction, as with addition, the error in the answer is equal to the sum of the errors in the individual amounts, *i.e.*, $\pm$ 1,000 and $\pm$ 500 = $\pm$ 1,500.

When *multiplying* two rounded values together, it is usual to show both the result and the maximum possible error which could have occurred in it. This is derived as follows. The product of two values, 12,500 and 400, is 5 million. Assume that the larger value has been rounded to the nearest 500 and the other value to the nearest 50. Thus the true values might have read 12,749 and 424 which when multiplied together give a product of approximately 5·43 million. The maximum error in the original

product of 5 million is thus 0·43 million which is equal to an error of 8½ per cent. Therefore, the answer to the multiplication given above would be written 5 million ±8½ per cent.

With *division* the same principle is applied. If the total 300,000 is divided by 500, the quotient is 600. If the dividend had been rounded to the nearest 10,000 and the divisor to the nearest 10 units, then the maximum error in the above quotient of 600 is derived when we divide the smallest possible divisor into the largest possible dividend, *i.e.*, 496 into 304,999, which gives 615. In other words, there is a maximum error of ±15 in a figure of 600, which is equal to ±2½ per cent.

It should not be necessary to memorise the various formulae which are sometimes evolved for these operations, usually remembered at the expense of the principles on which they are based. All that needs to be kept in mind is that the maximum error is always possible. Generally speaking, no amount of juggling with figures can increase the accuracy of the result if the original data are liable to error. The result of any calculation involving approximation can be no more accurate than the least accurate of the figures used in the calculation. Thus, in stating the final result, it may be advisable to give it to at least one significant figure less than found in the least accurate factor employed in the calculations. The same point should be borne in mind when calculating means and medians, etc., from grouped frequency distribution where the class-interval is large.

Use of Ratios

Earlier in the chapter the value of relatives as against absolute quantities for comparative purposes was mentioned. The main feature to be remembered is that a ratio or percentage expresses the variation in the data, irrespective of its actual or absolute size. Thus an expansion in a firm's turnover from £500,000 to £750,000 is the same in relation to the first value as is a rise from £10,000 to £15,000, *i.e.*, 50 per cent. rise on the base year, *i.e.*, the year in which the £500,000 and £10,000 were earned.

The main danger to avoid in expressing variations in terms of ratios or percentages is the use of two different bases in the comparison, or more generally, of failing to make clear on which base the change has been calculated. Thus, if in Year 1 profits were £25,000, and the chairman made the following statement:

'In Year 2 profits rose 10 per cent., the following year 25 per cent., and last year 33 per cent.', the shareholders might be forgiven if they arrived at two very different results:

	Year 1	Year 2	Year 3		Year 4
(i)	£25,000	£27,500	£34,375	and	£45,833
(ii)	£25,000	£27,500	£31,250		£33,333

The first line is calculated on the assumption that each percentage rise is based on the figure of the immediately preceding year; in the second line the percentages are all worked on Year 1 as base year. Whichever method is intended, it should be made clear which is to be the base year.

Such comparative ratios or percentages are a frequent source of confusion. If the two sets of quantities are widely different in absolute size, as in the examples of the sales given above, a mere percentage comparison may be quite misleading in so far as it may tell only half the story, or more seriously, it may suggest that the comparison made is justified when it most certainly is not. Thus, if a school teacher discussing the latest examination results with the head of a large coaching institution, states that 50 per cent. of his candidates obtained distinction in all subjects, and the coaching institution only 5 per cent.; can any conclusions be drawn? The answer is 'no'. It may be learnt later that 500 pupils of the coaching institution sat and only six at the teacher's school. Furthermore, it is highly improbable that a direct comparison can be made between the two teaching methods, until more is known of the calibre of the pupils, the amount of study done by the students, and numbers of staff. The first rule of statistics, 'compare like with like', has hardly been observed.

The averaging of percentages themselves requires care, where the percentages are each computed on different bases, *i.e.*, different quantities. The average is *not* derived by aggregating the percentages and dividing them. Instead of this, each percentage must first be multiplied by its base to bring out its relative significance to the other percentages and to the total. The sum of the resultant products is then divided by the sum of the base values as in (Col. 2) below, not merely the number of items.

Suppose the ratio of equity capital to total share capital for six public companies is as follows:

Company	A	B	C	D	E	F
%	100	50	50	40	50	25

The average ratio for all six companies is not the average of these ratios, *i.e.*, 315 ÷ 6, as set out in column (4) below, because this method makes no allowance for the different amounts of the companies' capital as shown in column (2) below. The correct method requires us to work from the actual figures of share capital and equity capital given in columns (2) and (3) below. Then each percentage is 'weighted', *i.e.* multiplied by the figure of total capital, their product added together as in column (5) and that total divided by the total weights. The correct answer is 43·3%, as compared with the answer of $\frac{315}{6}$ = 52·5% derived by using the incorrect method.

Percentage of Equity Capital in Six Public Companies

(1) Company	(2) Total Capital 000's £	(3) Equity Capital 000's £	(4) Ratio of Col. 3 to Col. 2 %	(5) Col. 2 × Col. 4 000's £
A	50	50	100	50
B	100	50	50	50
C	200	100	50	100
D	250	100	40	100
E	500	250	50	250
F	400	100	25	100
	1,500	650	315	15)650 = 43·3%

The reason for the difference is simple. In Column (4) there is an implied assumption that all the companies are of equal size (in terms of their issued capital) and consequently equal importance has been attached to the capital structure of the small and the large firms. In Column (5), however, correct importance has been given to each firm; *i.e.* the percentage of equity to total capital has been *weighted* in the ratio of their total capitals one to the other.

The same rules apply to the 'averaging' of averages, *e.g.*:

(1) Plant	(2) No. of operatives	(3) Average Weekly Output per operative	(4) Products in 00's
A	180	540	972
B	140	530	742
C	50	490	245
D	90	500	450
E	110	510	561
F	160	525	840
	730	3,095	730)3,810
			522 to nearest unit

The average output of all workers in the six plants is not the average of the six plant averages, *i.e.* 3,095 ÷ 6 = 516. This is wrong. The correct method is given in the last column; the products of multiplying Columns 2 and 3 together and dividing their sum by the total of operatives, *i.e.*, Column 2. The true average per operative is thus found to be 522 per week. The reason for this procedure, known as weighting, is that more importance or weight must be attached to the larger plants, *i.e.*, those with more operatives and the greater total output.

Good and Bad Statistics

For the layman statistics may often possess a significance far beyond their real importance. For example, a speech containing a large number of statistics is often regarded by audiences as much more impressive than one which concentrates on ideas and principles. Much the same is true of the printed figure, and as many a statistician knows, what was once a hopeful guess in the committee room can all too often later appear to haunt him in a published report. The latter will in all probability be extensively quoted with all the authority of the office from which the guess emanated! There are good statistics and bad statistics; it may be doubted if there are any perfect ones which are of any practical value. It is the statistician's function to discriminate between good and bad data, to decide when an informed estimate is justified and when it is not.

Poor statistics may be attributed to a number of causes. There are the mistakes which arise in the course of collecting the data, and there are those which occur when those data are being converted into manageable form for publication. Still later, mistakes arise because the conclusions drawn from the published data are wrong. The real trouble with errors which arise during the course of collecting the data is that they are the hardest to detect. It is virtually impossible to check whether an interviewer should have ticked *Yes* instead of *No* as the answer to a given question. Like the rest of mankind, interviewers make mistakes; they don't always ask the right questions and they sometimes write down the wrong answer.[1] When the questionnaires and schedules are returned to the Head Office for tabulation a new source of error appears. The answers may be incorrectly transferred from the schedules to the punched cards or tabulations; but good supervision can reduce this risk considerably. Sometimes, however, the answer given on the schedule has to be classified. This is at best an arbitrary procedure and mistakes in classification arise. Once the tables have been prepared detailing the results of the enquiry, their contents are analysed. Here too, a great deal more can be read into some statistics than the people who provided them ever dreamed of!

A weakness frequently encountered in reports which contain published statistics of trade, unemployment and other economic or social affairs, is the failure to consult with sufficient care the source from which a total or figure has been taken. Unemployment figures for Great Britain may be incorrectly related to population figures for England and Wales; pre-war employment figures in certain industries may be freely compared with current data yet there is no evidence that the author realises that owing to a re-classification of Ministry of Labour statistics, the two totals may cover very different fields. This is especially true of index numbers which are quoted to measure changes in quantities and values over periods of time. For example, changes in the present cost of living can be measured from month to month by the Index of Retail Prices, but this index cannot be directly compared with either the Interim index for the period 1947-52-56 or the pre-1947 Cost of Living index. In the second chapter, great emphasis was laid upon the need for verifying definitions

[1] Some of these problems are discussed in Chapter XIII.

and sources of data taken from published documents; no apology is made for returning to this theme because inaccurate extraction of published data is an extremely prevalent disease.

Conclusion

The object of the previous somewhat depressing paragraphs is simply to remind the student who has just begun to perform some simple calculations with numerical data that, however accurate his arithmetic may be, it cannot improve the data one iota. Poor statistics yield unsatisfactory results and as we plunge further into the technicalities of what is called statistical method, the reader should not allow himself to consider that statistical calculation is the end-all of statistics. Every set of figures reaching the statistician's desk should be scrutinised, their precise definitions learned and their source examined to consider what errors could have arisen. Only then has statistical analysis any purpose. The reason for harping on this point is that many of us forget these simple rules; or erroneously believe that a competent statistician collected the data and obviously would have checked them!

CHAPTER IX

THE BASIS OF SAMPLING

Introduction

Throughout the discussion of averages and measures of dispersion it has been assumed that summarisation and description of the various types of frequency distribution was the primary object of such exercises. Such an impression would be confirmed by the description of these statistics in the opening chapter as *descriptive* statistics. Most of the frequency distributions used to illustrate these statistics comprised a complete count of all the members or units in the 'population', used here in its statistical sense. Thus the data relating to output of factory operatives was based on a complete count of all employees in those works; Table 9, showing the distribution of the firms in the radio industry employing more than 10 workers, was also based upon a complete enumeration or census of the relevant population, just as was the distribution of incomes set out in Table 6. The size of the populations varied considerably from the 180 workpeople in Table 3 to the 23 million income earners in Table 6.

It is not often, however, that the statistician has access to data based upon an up-to-date census. Such enquiries are expensive in both time and money. The population census is held only at ten year intervals and the full scale census of production at three year intervals, although the Inland Revenue can produce detailed statistics such as those in Table 6 at yearly intervals, but with some delay. These enquiries are extremely important from the administrative point of view and it is to serve such needs that they are carried out. Any benefits which the statistician derives from access to such data are merely incidental to the main object. In brief, the statistician can very seldom hope for an enumeration of all the members of a population in which he is interested. Thus when the B.B.C. Audience Research unit wishes to know viewers' reactions to last night's TV programmes it cannot hope to ask all viewers for their opinions. It would be very costly and would take so much time that many viewers would have

forgotten all about the programme by the time they were asked for their opinions.

The solution to the problem is for the research unit to interview only a small proportion of the viewing public and on the basis of their findings to infer the views of the viewing public as a whole. Clearly, the group of viewers to be interviewed must be representative of all types of viewer, for example, both working class and middle class households must be included. Such a survey of a small proportion of the relevant population, in this case all households with B.B.C. television who actually watched the programme, is known as a *sample* survey. Exactly the same procedure is followed by the public opinion polls which measure the fluctuations in the public's support for the policies of the political parties. Experience shows that before a general election such polls can usually forecast the outcome quite accurately although they may be based upon a sample of less than one in 10,000 voters.

In themselves sample results or *sample statistics* are usually of little real value or interest. They are collected because the statistician is interested in the parent population from which the sample is drawn. On the evidence of such sample statistics it is possible to derive information about the population, such as estimating what the statistician calls the *population parameter*. For example, a sample mean provides an estimate of the population mean. Unless it were possible to generalise sample results in this way, sampling would be of negligible value to the statistician. At best they would provide some indication of the nature of the population from which the sample was drawn. But without modern statistical theory such an inference would be little better than a guess. The theory of sampling makes it possible not only for inferences and conclusions to be drawn from sample data but enables the statistician to make precise probability statements about the reliability of such conclusions. For example, let it be assumed that a sample of 100 adult British born males has been assembled and it is found that the mean weight of the sample is 150 lb., while the standard deviation of the sample is 15 lb. From this information the statistician could make the prediction that if further samples of equal size were taken, 95 times out of a hundred their means would lie between 147 and 153 lb. But because one or more sample means has a

value of 150 lb. can it be inferred that this value is equal to the value of the population parameter? Or, if it is regarded as an estimate of that value, how good an estimate is it?

Problems of Estimation

Since the population mean has a certain unique value, it follows that any estimate based upon sample data must be either right or wrong. Unfortunately, there is no way of knowing which it is. In practice, however, we speak of there being a certain probability that the population mean lies between certain limits. The 'certain probability' indicates the degree of confidence we have in our statement that, for example, the population mean lies between 147 and 153 lb. In this case we are 95 per cent. certain and in so stating our confidence we concede the chance that our prediction regarding the mean weight of all males will be wrong 5 in a hundred times. Sometimes the statistician will make his prediction with only one chance in a hundred of being wrong. In such cases he is described as being 99 per cent. certain or confident that the specified range of values covers the population mean.

It is worth while considering for a moment what we mean by being 'certain'. Each of us can be certain that he will not live 200 years. In this case we are 100 per cent. certain, but such confidence about the course of future events is quite exceptional. A person may be certain at 5 p.m. that at 5.30 p.m. he will leave his office and return home. He might, however, drop dead at 5.15. For most people the chances of such an occurrence are so slight that the possibility is ignored; but it undoubtedly exists. Thus there can be no absolute or 100 per cent. certainty that the person in question will leave the office at 5.30. On the other hand, leaving home next morning the weather may look ominous and he may take his raincoat. He is not certain that it will rain, but past experience of such portents as overcast skies tells him that it is more likely to rain than not. He therefore decides to take a coat. Most decisions are the outcome of weighing up such chances, although the mental processes may be so rapid that we are oblivious of their nature. Thus when the statistician affirms his belief in an estimate at the 95 per cent. level of confidence, he is doing what we all do every day of our lives. But whereas we say, 'well, it probably will rain', he indicates precisely just how

much confidence he has in his forecast, *i.e.*, 95 chances in 100 that it will rain, 5 in 100 that it will remain dry. Our reason tells us that given such odds, we should take a raincoat. In the same way the statistician will act on the evidence of his sample data.

From the layman's point of view a weakness of the estimate of the population mean lies in the fact that the statistician's estimate of a parameter is not precise. It is usually stated as lying between specified limits, sometimes called confidence limits. This criticism is valid but clearly much depends on the range of the limits. An estimate of the population mean such as '150 plus or minus 30 pounds would be useless for most purposes; equally, 150 lb. plus or minus 3 lb. is for many purposes almost certainly adequate. If it is necessary to narrow the range of values between which the parameter is believed to lie, then it can be done by taking a larger sample. It would be quite practicable to take a sample which would provide an estimate of the population mean within a quarter of a pound either way. This will probably satisfy even the most exacting requirements. The *precision* of the estimate, as it is termed, is not the result of chance. It is dictated by the needs of the survey and any specified requirements can be met even if at extra cost.

A brief recapitulation of this chapter so far will probably help the reader. Three points have been made. The first was that the primary purpose of sampling was to derive statistics which would yield information about the population. For example, if we want to know what proportion of the public drinks coffee for breakfast, we can ask such a question of a representative group of the public and on the evidence of the sample statistic estimate the population parameter, *i.e.*, the proportion of the public which does drink coffee. The second point is that the statistician can never give an exact figure as his estimate of a parameter; he states the limits within which he believes it to lie. Since the precision of sample statistics can be determined by the statistician, this apparent weakness is relatively unimportant. Finally, the statistician can *never* be certain of the correctness of his estimate of the parameter or any other conclusion based on sample data.[1] Any inference or conclusion based upon sample data is made at a specified level of confidence, usually at the

[1] Some further statistical tests based upon sampling theory are discussed in the next chapter.

'5 per cent. level'. This signifies the statistician is confident that in the long run he will be right 95 times out of every 100. Since such odds imply a very considerable confidence in the correctness of the inference and if there are no other data to suggest alternative conclusions, it remains merely to act on the inference. The next section of this chapter is devoted to an explanation of the theory underlying sampling and confidence levels.

The Normal Curve

Everybody has some idea of the meaning of the term 'probability' but there is no agreement among scientists on a precise definition of the term for the purpose of scientific methodology. It is sufficient for our purpose, however, if the concept is interpreted in terms of relative frequency, or more simply, how many times a particular event is likely to occur in a large population. When we say the probability of obtaining 'heads' as a result of tossing a coin is equal to ½, we mean that if the coin were tossed a large number of times the proportion of heads to be expected is one half. Similarly, if a set of ten coins were tossed simultaneously we should expect to get five heads rather than any other number. If we continued to toss the ten coins it would not be a matter for surprise if 3, 4, 6 or 7 heads were recorded quite frequently and from time to time we might even get all the coins falling heads and on other occasions all tails. But in the long run, assuming all the coins to be true, we should expect to record 5 heads more frequently than any other score.

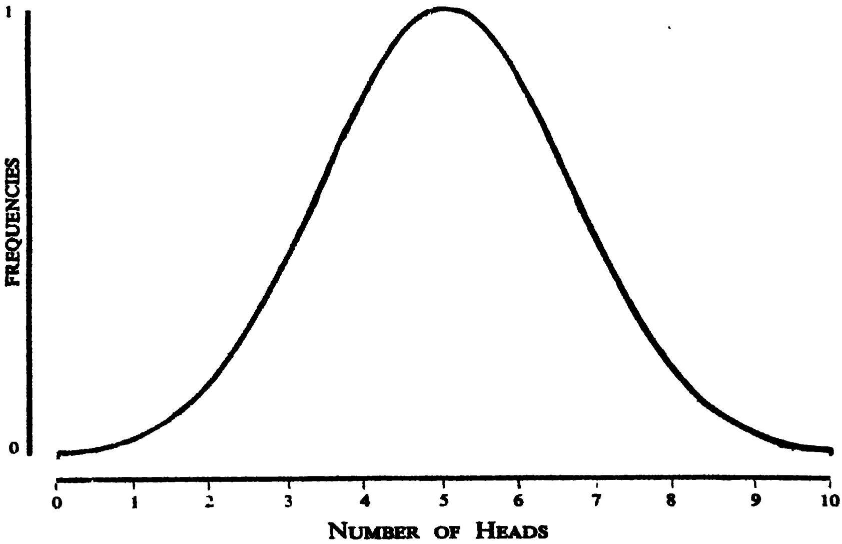
It can be demonstrated mathematically that where there are only two alternative outcomes to an event, *e.g.*, heads or tails with a spun coin, and the coin is so tossed that chance alone determines which way it falls, then if the experiment is repeated a very large number of times, the distribution of results can be predicted. Such a distribution is known as the *binomial* distribution. Take, for example, the experiment of tossing ten coins simultaneously and noting the number of heads. In the early stages one head might follow six heads and then three heads; no particular order would be apparent. But as the number of tosses grew, so the distribution of frequencies of particular numbers of heads would take on a definite pattern. Such a frequency distribution with the values 0–10 heads marked off along the base and the corresponding proportionate frequencies

marked off along the vertical axis, would resemble a histogram similar to that depicted in Figure 12 on page 81.

The same experiment could be performed tossing sets of 100 coins or even a thousand coins at a time. With such large values to be plotted and if the experiment were continued for a long time the resultant histogram would have an outline resembling tiny serrations rather than the clearly defined steps of the histogram in Figure 12. In fact, for very large numbers of tosses the histogram would cease to resemble that histogram but would approximate to a smooth curve similar to that portrayed in Figure 16. This curve is bell shaped and symmetrical. These characteristics reflect the fact that the maximum frequencies are recorded by the mean value of the distribution, *e.g.*, five heads, while other values which deviate by only a small amount from the mean also occur quite frequently. But it is very evident that as the deviation between any recorded value and the mean increases, so the frequency of that value declines. There are very few cases indeed where all the coins came down heads and there are equally few where they were all tails.

While it may not be immediately apparent, the reader will on

Figure 16



reflection realise that this curve portrays the distribution of a certain sample statistic. Each set of ten (100 or 1,000) coins tossed is merely a sample of an infinitely large population of such tosses. The mean number of heads recorded in each sample is an estimate of the proportion of heads we are likely to get if every coin in existence could be tossed, *i.e.*, 5 out of 10, implying that the chances of heads or tails with a spun coin are in fact 50:50. If instead of tossing coins a series of samples, each consisting of 50 adult males of British birth, had been taken and for each sample the mean height was obtained, then the distribution of sample means would approximate closely to the curve depicted in Figure 16. Such a distribution is known as a *sampling distribution*, *i.e.*, a distribution of sample statistics such as sample means, proportions, standard deviations, etc. The curve based upon this type of distribution is known as the Normal curve.

The *Normal curve* is of fundamental importance to statistical theory. While it is always bell-shaped and symmetrical about its mean, its actual shape is determined by the standard deviation of the distribution. For example, two Normal distributions can have the same mean, but if for one the standard deviation is much larger than that for the other, then the curve of the former

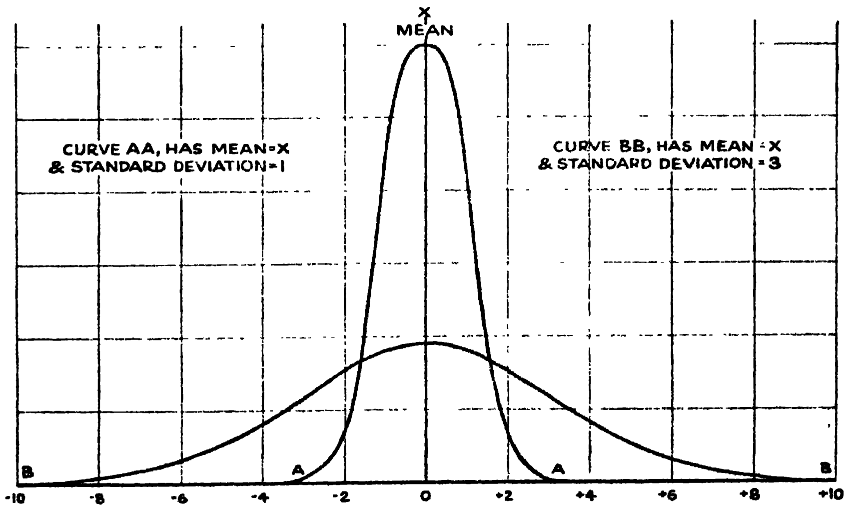
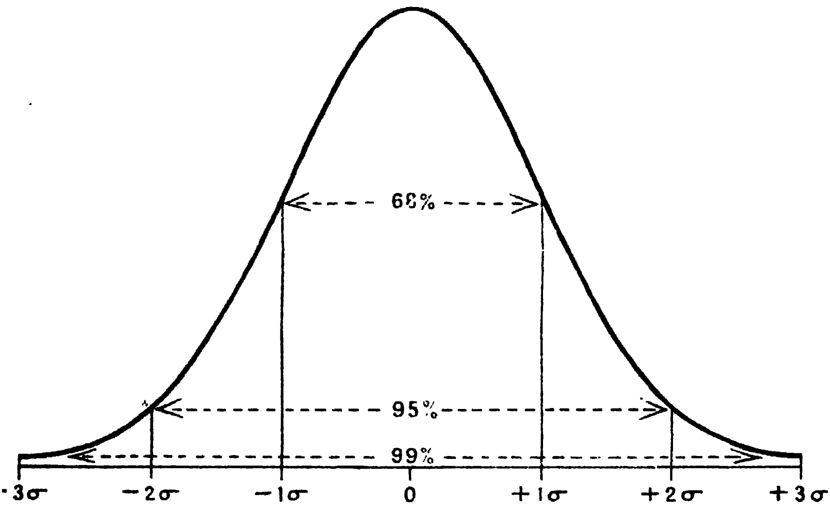


FIG. 17 NORMAL CURVES WITH THE SAME MEAN BUT DIFFERING STANDARD DEVIATIONS

distribution will be flattish and for the other more peaked as in Figure 17. In other words, the Normal distribution can resemble a tall single-peaked narrow curve or a flattish long-tailed broad based curve. It is, however, always unimodal, *i.e.*, a single peak, and symmetrical about that peak. The Normal curve has one other highly important characteristic. No matter what its shape, *i.e.*, broad or narrow, the area beneath the curve is distributed in a particular way. ~~From Figure 18~~ it can be seen that if vertical lines termed 'ordinates' are drawn from the base to the curve at intervals of one standard deviation from the mean ordinate, the proportion of the area under the curve bounded by the ordinates drawn at one, two and three standard deviations on either side of the mean is approximately 68, 95 and 99·7 per cent. respectively.

Figure 18

DISTRIBUTION OF AREA UNDER NORMAL CURVE.



So important is the Normal distribution in statistical theory that specially prepared tables give the proportion of the area under the curve enclosed between the mean and other ordinate. Since the Normal curve is symmetrical, it is sufficient to state the areas in this way, *i.e.*, for one side of the curve only, rather than as has been shown in Figure 18 above, where two ordinates equi-distant from the mean ordinate have been drawn on each

side of it. To derive the area enclosed in such a case it is merely necessary to double the figure shown in the table of areas. The reason for setting the table out in this way is that sometimes the statistician is only concerned with one side of the Normal distribution. The following table merely reproduces a few important values taken from the complete table, but it will serve to show what sort of information such a table provide~~s~~.[1] The distance along the base between the mean ordinate and the other is always measured in terms of the standard deviation of the distribution and ~~is written x~~/σ. For most purposes the only values of x/σ that interest the statistician are those between 1·96 and 3·09 and of those it is the 1·96 and 2·58 values which are customarily used in practical work. The reasons for concentrating on these few values will become apparent to the student during the next few pages.

x/σ	1·96	2·33	2·58	3·09
Area	0·4750	0·4900	0·4950	0·4995

If ordinates are drawn on *both* sides of the mean at intervals corresponding to the above values of x/σ, then the percentages of the total curve enclosed between those ordinates are 95, 98, 99 and 99·9 respectively. It will be seen that these percentage figures are derived by doubling the figures in the lower row of the table above and expressing the result in percentage terms.

Now, instead of visualising the Normal curve as an area divided into parts, consider it as a large collection of sample statistics, *e.g.*, a very large number of sample means. Clearly, the mean of all the sample means is the best estimate we have of the population parameter, *i.e.*, the mean of that population from which the samples have been draw~~n.~~ It is apparent that within a range of 1·96 standard deviations about the mean of the Normal distribution 95 per cent. of all sample means are to be found. Similarly, 99 per cent. of the sample means lie within a range of 2·58 standard deviations about the population mean, while within 3·09 standard deviations about the mean we find 99·9 per cent. of the sample means. In practice, however, there is no time to take such a large number of samples from any

[1] Many of the more advanced texts reproduce the entire table of areas of the Normal or probability curve in an appendix.

particular population, we have to be satisfied usually with a single sample. Fortunately, because the sampling distribution of the statistic is known, the chances of a single sample statistic deviating by a given amount from the mean of such a distribution, *i.e.*, the population parameter, are known. For example, the probability that a single sample statistic will deviate by more than 1·96 standard deviations from the population parameter, *i.e.*, the mean of the sampling distribution is only 5 in 100. We can therefore be confident about estimating, on the evidence of a single sample statistic, the value of the corresponding population parameter. Since 95 per cent. of all sample means will lie within a range of 1·96 standard deviations about the mean of the distribution of sampling means, it follows that the chances of a sample mean based on a single sample deviating by more than 1·96 standard errors from the population mean are only 5 in 100, or 1 in 20. Such then is the basis of the statement made earlier (page 152) that the statistician could assert with 95 per cent. confidence that the mean weight of all adult males, *i.e.*, the population of such males, was between 147–153 pounds

As implied in the previous sentence, and as was shown above in the excerpt from the table showing areas under the Normal curve, the proportion of the means lying between 1·96 standard deviations about the mean is 95 per cent. In practice, to give our results at the '5 per cent. level', the standard error is usually multiplied by 2. The calculation is so much more rapid and it also increases the confidence limits to give us just that little more confidence in our estimate. For these reasons we shall in the next sections use a range of 2 standard errors to denote the 5 per cent. level. We are, after all, here concerned with matters of principle, not arithmetic.

Standard Error of a Statistic

For any sampling distribution we can calculate the mean and standard deviation. The variation between large numbers of sample means, which produces the dispersion about the mean of the sampling distribution, arises from what are termed random sampling errors. The 'standard deviation' of such a sampling distribution is known as the *standard error*. This is computed for any sample statistic and is a measure of the precision of that

statistic as an estimate of the corresponding population parameter. The standard error of a sample mean is given by the formula s.d./$\sqrt{n}$. In this fraction the numerator is the standard deviation calculated from the sample. The formula really requires the standard deviation of the population, but since this is unknown we substitute the sample standard deviation. Fortunately it can be shown mathematically that where the sample is large, the inaccuracy introduced by the substitution is unimportant.

From the above formula it is apparent that the standard error of the mean is dependent upon two factors, the variability in the population and the size of the sample. The point has already been made that successive samples may produce different statistics and the variation between sample statistics is dependent upon the distribution of the characteristic within the population. For example, if the weights of all adult males were within a range of 10 lb. about a mean of 150 lb., then clearly no sample mean could be below 140 or above 160 lb. Since, however, the actual weights range from about 100 to 200 lb. or more, it follows that sample means of any value between these two limits are possible. Thus the standard error of the sample mean must be dependent in part on the variability within the population which is measured by the standard deviation.

Increased precision in the sample estimate of a parameter can be obtained by increasing the size of the sample. There is a definite relationship between the standard error and the size of the sample but as is apparent from the above formula for the standard error of the mean, it is not a direct relationship. Since the formula uses the square root of the sample size, it follows that in order to halve the standard error of the mean, the sample size must be increased fourfold. It is for this reason that given the standard error of a statistic based on a large sample, an increase in its precision can be achieved only by considerably increasing the size of the sample. Since this is an expensive procedure, the statistician endeavours to select that size of sample which will give the maximum precision in the sample estimate consistent with a given outlay of funds. In other words, the statistician decides first upon the level of precision he requires, *i.e.*, the standard error he is prepared to accept, and then he calculates the size of the sample required.

The above points can be simply illustrated. Assume a sample of 100 adult women with a mean height of 63 inches and a standard deviation of 2 inches. The standard error of the mean is obtained by substituting the appropriate values in the formula s.d.$/\sqrt{n} = 2/\sqrt{100}$ equals 1/5th inch. It can be inferred from these data that the assertion that the population mean lies within the range of 63 inches + or — 2/5ths inches is correct at the 5 per cent. level. Suppose that greater precision is required in the estimate of the parameter, *e.g.*, the standard error should be one tenth of an inch, which is half the present error. Substituting in the equation we get $1/10\text{th} = 2/\sqrt{x}$ which yields a value of 400 for x. Thus, to reduce the standard error by half the sample had to be quadrupled.

There is an alternative available, albeit a poor one, if the sample cannot be increased. The population mean could be assumed to lie within a range of 1/10th inch about the sample mean of 63 inches but the chances of this being true are only 68 in 100. In other words, the statistician can improve the apparent precision of his estimate by accepting a greater risk of being wrong. With a range of only one standard error about a statistic he runs the risk of being wrong 32 in 100 times, *i.e.*, virtually 1 in 3 and for all practical purposes this risk is too high. Convention dictates that most estimates be given at the 5 per cent. level, *i.e.*, where the chances of being wrong are only one in twenty. With a given sized sample, the statistician must choose between the precision of his estimate and the degree of confidence he has in the result.

Just as it is possible to estimate the population mean from sample data at a given level of confidence, so an estimate of a population *proportion* from a sample data can be made. For example, a sample of 900 electors reveals that 45 per cent. intend to vote Conservative in the forthcoming election. What is the probable proportion of the total electorate which will vote Conservative? First, we must derive the standard error of the proportion which, like that of the mean, is based upon the distribution of sample proportions and is given by the formula $\sqrt{\frac{p \times q}{n}}$ The symbols p and q represent the proportions of the sample possessing or not possessing the relevant characteristic, in this case the intention to vote Conservative. With a coin, the odds of

obtaining a head were one half and those for tails also one half, so that certainty was represented by $\frac{1}{2} + \frac{1}{2} = 1$. In the same way, so $p + q$ equal 1 and the value of q is derived from the difference between 1 and the value of p. In this case $1 - p$ equals 0·55. The formula for the standard error of the proportion after substitution of the symbols reads

$$\text{s.e. }\% \sqrt{\frac{(0{\cdot}45) \times (0{\cdot}55)}{900}} = {\cdot}017$$

The calculation gives a result of approximately 1·7 per cent. As with sample means, so the distribution of sample proportions is such that 95 per cent. of a distribution of sample proportions will lie within a range of two standard errors about the population percentage. Given this knowledge it may be inferred from the above data that the proportion of potential Conservative voters in the electorate is $45 \pm 2(1{\cdot}7)$ per cent. or between 41 6 and 48·4 per cent., the estimate being made at the 5 per cent. level Using the above formula the student can now calculate for himself the size of sample required to give a result within the limits 44 to 46 per cent.[1]

Summary

Sampling is fundamental to all statistical analysis. Sample statistics are merely estimates of the value of population parameters and individual sample statistics will differ from that value by what are known as random sampling errors. Mathematical theory demonstrates that the distribution of such random sampling errors tends to be Normal and observation supports the theory. Knowledge of this distribution in respect of sampling variation enables the statistician to make rigorous inferences concerning the population on the basis of sample data. A mistake which is commonly made is to assume that it is the population which is normally distributed. This is not the case; very few populations are so distributed. Generally speaking, both the distribution and size of the population are irrelevant in sampling; it is only the statistics derived from successive samples that in the long run distribute themselves normally.

It cannot be too strongly emphasised that conclusions based on the evidence of sample statistics are valid only if the sample

[1] The answer is 9,900. The student unable to agree this answer should turn to page 198.

has been selected in such a way that every unit in the population has a known and equal chance of selection. Such samples are *random* samples, and even if the adjective is later omitted in the text the implication of the term 'sample' as used by the statistician is that it has been selected by methods which ensure randomness. Some of the more generally employed methods of choosing a sample are discussed in Chapter XI.

CHAPTER X

SIGNIFICANCE TESTS

In the last chapter it was explained that even if nothing is known about a population it is possible on the basis of a random sample from that population to derive reliable information about its nature. This important statistical technique is possible because of our knowledge of what are termed the 'sampling distributions' of the various sample statistics. The 'standard error' of a sampling distribution, for example sample means, indicates the range of deviation from the population mean which can be expected with a given frequency in the means of a large number of random samples. Since it is known that the proportion of a large number of sample statistics lying within two standard errors of the population parameter is 95/100, the chance of a deviation between the sample statistic and the population parameter greater than twice the standard error of the statistic is approximately only one in twenty. In practice, this means that when a solitary sample is drawn, and this is all that is usually practicable the statistician can be reasonably confident that it is representative of the population and not a collection of the extreme items in that population. He can therefore feel reasonably certain that any estimate of the population parameter based on that sample statistic is likely to be accurate.

This knowledge of sampling distributions is not only used in problems of *estimation*, as they are termed. There is a second group of problems which have to be resolved on the basis of sample results; these are known as the *testing of hypotheses.*

Assume for example that a manufacturer believes that housewives in the North of England have a stronger preference for his product than the housewives in the South; or a doctor believes that a particular drug would lead to a more rapid recovery by the patient from a particular disease. Such beliefs as these are not founded on thin air; both the manufacturer and the doctor in this illustration have presumably grounds for their beliefs. The manufacturer may have just completed a tour of the country and

returned to head office with these personal impressions; the doctor after having treated a number of patients suffering from this particular disease with a variety of drugs has ascertained that in his experience at least one drug is superior to all others. The issue posed by these examples is whether or not the statistician can enable the manufacturer and physician to demonstrate that their beliefs or theories, or what the statistician terms 'hypotheses', are valid. For example, if another doctor treated patients with this particular drug, would he also find that it was superior to others? In other words, the doctor and the manufacturer are both basing their conclusions on limited evidence. Is it possible to show that what is valid for a sample of one doctor's patients is true for all others? The statistician can help in resolving such problems by using what are known as *significance tests*. As with the problems of estimation which were discussed in the latter part of the previous chapter, the statistician can never be absolutely certain that the results of such significance tests are valid. All he can do is to estimate the probability that the hypothesis is 'true' in the light of the information available. Using such techniques, an hypothesis may be tested at the 5 or 1 per cent levels (or any other!) and on the result we may be 95 or 99 per cent certain of the validity of our hypothesis. These ideas can now be illustrated by a simple example.

S.E. of Differences between Means

Assume that the mean consumption of beef in the North of England as given by a sample enquiry among 845 households is 50 lb. per annum with a standard deviation of 13 lb., while a similar survey in the Home Counties covering 1,440 households yields a mean consumption of 48 lb. and a standard deviation of 12 lb. On the evidence of these data can it be inferred that beef consumption is greater in the North than in the Home Counties? In the language of statistics, is there a statistically *significant* difference between Northern and Southern households, or is it more likely that the difference of 2 lb. between the sample means can be attributed to chance, *i.e.*, random sampling errors? When a difference such as this cannot reasonably be attributed to sampling errors, we say it is significant. Since, as we have seen, we can never be 100% certain about inferences drawn from sample statistics, we have to state at what level we consider the

difference to be significant, *e.g.*, either at the 95% or 99% level of confidence.

The question posed above can be phrased as follows: 'Can it be assumed that these two samples were drawn from the same population?' The statistician answers this question by employing a 'significance test', the basis of which is the 'hypothesis' that the difference between the sample statistics can be entirely explained by sampling variation and there is no real or *significant* difference between the sample means. More simply, it implies that the two samples are drawn from the same population. This is known in statistics as the *null hypothesis*. The result of the significance test will either strengthen or weaken confidence in the hypothesis; it can never prove *conclusively* that it is false. Thus we shall either *accept* or *reject* the hypothesis at a given level of confidence.

Our knowledge of sampling distributions is again invoked, and in this case we want to know what the chances are that a difference of 2 lb. might occur between two random samples drawn from the same population. To answer this question we first calculate the standard error of the difference between means, using the formula:

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Substituting the known values of the n and σ^2 in the formula we get:[1]

$$\sqrt{\frac{13^2}{845} + \frac{12^2}{1{,}440}} = \sqrt{0{\cdot}2 + 0{\cdot}1} \quad \sqrt{0{\cdot}3}$$

which yields an answer of 0·547 lb. This is the standard error of the difference. From our knowledge of the behaviour of sample means it follows that the chances of a difference between means greater than twice the standard error is less than 5/100 and that a difference more than two and a half times as great is likely to arise by chance only once in a hundred times. The observed difference between the sample means is 2 lb. which is almost four times as great as the standard error. Such a difference could arise by chance only once in about ten thousand times. In other words, the odds against the likelihood that the two samples have been drawn from a single population are so small that the null hypothesis may reasonably be rejected. The observed difference

[1] Note that the sub-scripts 1 and 2 to n and σ^2 merely serve to distinguish the two samples.

is *statistically significant* and it can be assumed, in this case with considerable confidence, that the average northern household eats more beef than its southern counterpart.

S.e. of Difference between Proportions

The same ideas can be illustrated by an example involving a difference between proportions. Assume that a market research agency learns from a survey of 1,000 households that 30 per cent of them regularly purchase a particular branded produce. A few months later the same agency carries out a similar survey and learns that 34 per cent of the households sampled purchase this product. Can it be concluded that there has been a four points increase in the product's popularity between the two surveys, *i.e.*, a significant increase, or is it possible that the difference can be attributed to sampling variation? The standard error of the difference between proportions is given by the following formula:

$$\text{s.e.\%} = \sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}$$

and substituting the observed values for p, q, and n we get

$$\sqrt{\frac{0{\cdot}3 \times 0{\cdot}7}{1000} + \frac{0{\cdot}34 \times 0{\cdot}66}{1000}} = {\cdot}021$$

which yields a result of 2·1 per cent. The observed difference of 4 per cent between the proportions is not quite twice as large as the standard error. We conclude, therefore, that the probability of random sampling errors accounting for the difference between the two percentages is just over 5/100. In other words, the statistician will be rather less than 95 per cent. confident that the change in the percentages is significant, *i.e.*, that it reflects an increase in popularity of the product. In such circumstances the statistician will probably err on the side of caution and accept the null hypothesis, *i.e.*, the difference is not significant.

It is of the utmost importance to realise that such significance tests can never finally or absolutely prove or disprove a hypothesis, they merely 'prove' beyond all reasonable doubt. If the statistician rejects the null hypothesis, he does so because the evidence adduced by the significance test is such as to weaken his confidence in it to such an extent that any other conclusion

would be unreasonable. For example, in the first illustration of the difference between means, where it emerged that the observed difference was only likely to arise by chance once in about ten thousand samples, the statistician would have no hesitation in rejecting the null hypothesis. Unfortunately, the results achieved by such tests are often similar to that obtained in the second illustration, *i.e.*, close to the level of confidence at which it is decided that a result is significant. It must be emphasised that there is nothing sacrosanct about the 5 per cent level any more than there is about the 1 per cent level. Experience has shown statisticians that the 5 per cent level is in the long run adequate for most of their needs; but if the experiments so require, then the significance of the results can be tested at a higher level. For example, a steelmaker selling a special alloyed steel, upon the toughness of which lives might depend, will want to know that the quality of his product is consistently above a certain level. By special significance tests based on sampling procedures he will ensure that the chances of sending defective goods will not be greater than, say, 1 in 1,000. A manufacturer of a consumer article may have a similar system of quality control and be well satisfied to know that the chances of despatching defective goods is not more than 1 in 20. The former will test at the ·01 per cent level, the latter at the 5 per cent level.

The application of such significance tests can be illustrated by a further simple example. Let us assume that over a long period a pottery factory has been accustomed to an average rate of defective products from the kilns of 10 per cent. A new firing technique is demonstrated and preliminary tests show that in 100 samples the rate of loss from this method is 7 per cent. The manager of the works believes that this improvement can be directly attributed to the new firing technique and is anxious to adopt the new firing methods. The board of the company is sceptical. Can the statistician help to resolve this problem? The standard of error of a proportion, it will be recalled, is given by the formula $\sqrt{\frac{pq}{n}}$ where n is the sample size, p the proportion of the sample possessing the characteristic and q the balance of the sample. The statistician must test the hypothesis that the variations in the proportion of defectives in samples

equal to the observed difference of 3 per cent can quite often arise by chance. The tests will show us what in fact are the chances of such a difference being encountered. Substituting the known values in the formula $\sqrt{\frac{0{\cdot}07 \times 0{\cdot}93}{100}}$ we get a standard error of 2·7 per cent. The observed difference between the proportions of defectives is only 3 points, which is little more than the standard error. It can be concluded that the chances of a loss rate as low as 7 per cent from the old firing technique are quite high, *i.e.* this loss rate can easily arrive by chance. While further tests should be carried out with the new method, the available evidence strongly supports the argument that the new firing technique will not significantly reduce the proportion of defectives.

The same test can be illustrated by another example. A random sample of 2,000 households possessing television sets capable of receiving both B.B.C. and I.T.V. programmes reveals that 1,050 declare a preference for the latter programmes and 950 prefer the B.B.C. May it be concluded from the survey that a majority of all households which can receive both programmes prefer I.T.V.? There are two ways of answering this question. We can assume that the population of households are evenly balanced in their preferences and that the difference revealed by the sample is due to sampling errors. If this is the case, then the proportion preferring I.T.V. is 1,000/2,000 or 50 per cent. We can calculate the chances that in two samples of 2,000 households drawn from a population in which the proportion with a given characteristic is 50% we might get sample proportions as divergent as 950/2,000 and 1,050/2,000, *i.e.*, 47½ and 52½%. The standard error of a proportion is given by the formula $\sqrt{\frac{pq}{n}}$ which, when the above values are substituted, reads $\sqrt{\frac{50 \times 50}{2{,}000}}$ The result is equal to approximately 1·1 per cent. If a large number of random samples of 2,000 households were taken, we may therefore expect 95/100 sample proportions ranging between $50 \pm 2.(1{\cdot}1)\%$. The actual observed sample proportions of 47½ and 52½% are outside the limits set by the test at the 5 per cent level, but within the limits of $50 \pm 2{\cdot}58$

(1·1)%. In other words, the difference between the samples is significant at the 5 per cent level, but it is not significant at the 1 per cent level. If a new advertising policy involving substantial expenditure on T.V. advertising, etc., is contemplated it might be advisable to take further samples before a decision is taken.

Chi-square test

An alternative method of solving this last problem is to use another test known as the *chi-square* (pronounced ki) test. This is written as follows χ^2 the name being derived from the Greek letter '*ki*'. This test is used when an observed distribution of frequencies must be compared with the expected distribution on the basis of some hypothesis. For example, in the above case, if we set up the hypothesis that the programmes are equally popular, the expected distribution is 1,000 households to each programme; the observed distribution is 1,050 and 950. The formula for chi square is written $\chi^2 = \Sigma\frac{(O - E)^2}{E}$. The data and calculations can be presented in the form of a table set out hereunder:

	Observed	Expected	Differences	Differences^2
I.T.V. ..	1,050	1,000	+50	2,500
B.B.C. ..	950	1,000	−50	2,500

The formula requires us to derive for each cell the difference between the observed (O) and the expected (E) frequencies and then square them. In this case the differences are 50 and when squared they become 2,500. These figures are then divided by the expected frequency for that particular sub-group. In each case we get 2,500/1,000 which gives 2·5 in each of the two cells and these, as the formula indicates by the summation sign (Σ), are then added together, *i.e.*, $2{\cdot}5 + 2{\cdot}5 = 5$. This is the value of χ^2. We now compare this value with the value of χ^2 which for this particular number of sub-groups is given in specially prepared tables. We learn there that the relevant value of χ^2 is 3·84 at the 5 per cent level and 6·64 at the 1 per cent level. In other words, our calculated value of 5 is above the value which could be expected to arise by chance 95/100 times, so that our hypothesis that the difference can be attributed to chance is rejected at the 5

per cent level, but not at the 1 per cent, since values of χ^2 up to 6·64 could be expected with such samples about once in one hundred times. The student will note that the result of this test confirms the result derived by calculating the standard error of a proportion, given above.

The chi-square test is a highly important and useful statistical test of significance. When we were testing differences between means and proportions we used tests based upon the standard errors of the corresponding sample statistics. These were all based upon the fact that we could assume that their distribution was Normal. Actually, the Normal distribution, important though it is in statistical theory, is only one of several distributions encountered by the statistician. The distribution of χ^2, for example is a special one but, as with the Normal distribution, special tables have been worked out which we can employ to interpret the values of χ^2 even if we cannot demonstrate here the mathematical basis of the distribution. The value of the χ^2 test lies in its application to the comparison of several frequency distributions. We set up the hypothesis that the distributions of frequencies are believed to be the same for both or all distributions (there may be several), but because of sampling errors, the distributions appear to be different. Using this test, we can establish whether or not the differences between what are called the observed and expected frequencies may be attributed to chance, *i.e.*, sampling errors, or whether they are *significantly* different. The following example may help to explain both the purpose of the test more clearly, as well as demonstrating the method of calculating the value of χ^2.

T.V. Owners Classified by Income of Head of Household

	Under £750	£750–2,000	Over £2,000	Total
I.T.V. ..	320	160	20	500
B.B.C. ..	280	190	30	500
	600	350	50	1,000

The above data show the distribution of two samples of 500 viewers apiece, one group having declared a marked preference for the B.B.C. programmes and the other for I.T.V. Each sample is

classified by reference to the income of the head of the household. For example, those over £2,000 may be regarded as 'upper' class, those with £750–2,000 as 'middle' class, while those earning less than £750 per annum we may define as 'working' class. At first sight the larger number of working class households revealing a preference for I.T.V. might be construed as signifying that the lower income groups as a whole prefer I.T.V., and the middle and upper classes B.B.C. programmes. We assume for purposes of our test that the two samples have been drawn from the same population and that the difference between the distributions of frequencies is entirely the result of random sampling errors. We set up, in other words, the null hypothesis, *i.e.*, that there is no difference in the distribution by income groups of households preferring I.T.V. or B.B.C. and that the observed difference between these two samples can be explained by random sampling errors. Given this hypothesis we assume in effect that the two samples are taken from the same population so the best estimate we can get of the distribution of the population among these three classes is obtained by aggregating the samples. Having done this we can say that in the absence of sampling variation the distribution of both samples of 500 between the three income classes would be 300, 175 and 25 as set out hereunder:

EXPECTED DISTRIBUTION

	Under £750	£750-2,000	Over £2,000
I.T.V. ..	300	175	25
B.B.C. ..	300	175	25
	600	350	50

The next part of the calculation gives the differences between the observed and expected frequencies within each of the six cells:

I.T.V. ..	− 20	+ 15	+ 5
B.B.C. ..	+ 20	− 15	− 5

These differences are squared and then divided each by the expected frequency in that particular cell. The reader will see that the upper line of differences is the same as the lower except for sign, but when they are each squared the signs all become positive. In this case there is therefore no need to sum all six

values of $\frac{(O - E)^2}{E}$ individually, just the top three and then double their aggregate.

$$\chi^2 = 2\left[\left(\frac{20^2}{300}\right) + \left(\frac{15^2}{175}\right) + \left(\frac{5^2}{25}\right)\right]$$

$$= 2\left[\frac{400}{300} + \frac{225}{175} + \frac{25}{25}\right]$$

$$= 2[1{\cdot}333 + 1{\cdot}286 + 1{\cdot}0]$$

$$\chi^2 = 2(3{\cdot}619) = 7{\cdot}24$$

The value of χ^2 comes to 7·24. Reference to the table of values of χ^2, an excerpt from which is reproduced below, shows that when $P = {\cdot}05$ and $n = 2$ (explained below), *i.e.*, at the 5 per cent level, the value of χ^2 is equal to 5·99. This signifies that a value for χ^2 could arise by chance from samples of this size as often as 5 in a 100 times. The value of χ^2 when $P = {\cdot}01$ and $n = 2$ is 9·21. Since the calculated value of χ^2 for the above data is 7·24 it follows that such a value could occur more frequently than one in a hundred times as a result of sampling variation but not as often as one in twenty. On the evidence of this test we are forced to the conclusion that the null hypothesis is rejected at the 5 per cent level, but not at the 1 per cent level. There is, therefore, on the evidence of sample data and the significance test, sufficient reason for believing that the distribution of households according to income is different for those who prefer I.T.V. from those who prefer B.B.C. programmes.

The greater the value of χ^2, the more likely it becomes that the differences between the observed and expected frequencies cannot be attributed to chance. The student may already have noted that the value of χ^2 depends on two factors. The first is the actual size of the frequencies, or more precisely the differences, but these are of course directly related to the absolute size of the frequencies. The second is the number of cells. The more cells we have, the larger must be the aggregate of $\frac{(O - E)^2}{E}$ since the squared difference is always positive. The tables of the values of χ^2 take account of these considerations, especially the question of the number of cells. The following is a small part of a table of values of χ^2:

DISTRIBUTION OF CHI-SQUARE[1]

Degrees of Freedom (n)	Probability level: P			
	0·05	0·02	0·01	0·001
1	3·84	5·41	6·64	10·83
2	5·99	7·82	9·21	13·82
3	7·82	9·84	11·34	16·27
4	9·49	11·67	13·28	18·46

The column on the left contains the values of n, which is dependent on the number of cells. The row across the top shows the various levels of probability to which the values of χ^2 for given values of n relates. For example, the column headed 0·05 gives the values of χ^2 if we are testing our hypothesis at the 5 per cent level, 0·01 values for the 1 per cent level etc. When $n = 2$, a computed value of χ^2 must exceed 5·99 if it is to be regarded as significant at the 5 per cent level; if it is as much as 13·82, then such a result signifies that the null hypothesis can be confidently rejected, there being only 1 in 1,000 chances of it being in fact true. The symbol n refers to what the statistician calls the number of 'degrees of freedom'. This is a very important concept but at this level of statistical knowledge it is not necessary to understand its precise meaning. All that is necessary is to be able to check a computed value of χ^2 against the tables. But for this purpose we must know how to obtain the appropriate value of n. It can easily be derived by the simple equation $n = (c - 1)(r - 1)$ when c represents the number of vertical columns of data and r the horizontal rows. In the above example, $c = 3$ because there were three classes of households and $r = 2$ since there were two main groups of viewers. Thus $(c - 1) = 2$ and $(r - 1) = 1$; therefore, $n = 2 \times 1 = 2$.

Conclusions

Since nearly all statistical analysis is based on sample data, it will be apparent that significance tests form a most important branch of statistical technique. The few examples given in this chapter represent only the more important tests of general application. It is of the utmost importance that the student should realise that significance tests do not prove or disprove a hypothesis. As already explained, they merely increase or

[1] This table is abridged from Table III of Fisher: Statistical Methods for Research Workers, published by Oliver and Boyd Ltd, Edinburgh, by permission of the author and publishers.

weaken the statistician's confidence in his hypothesis, which by its very nature has been set up on the basis of other information available to him. The test is in a sense the last, not the first, stage of his research. Although the application of these tests may appear at first sight to entail no more than simple arithmetic, the interpretation of the results requires very considerable experience and, above all, knowledge of the data and the sample. The real purpose of this chapter is not so much to explain how the student can imitate the professional. It is first to explain something of the statistician's real work, and second to enable the reader to understand the published results of research in fields ranging from the social sciences to engineering and chemistry where these methods are continuously employed. One simple application of elementary sampling theory is provided by the use of Quality Control in industry, the principles of which are explained in Chapter XX.

CHAPTER XI

SAMPLING METHODS

Sampling has been described as the selection of a proportion of the population to obtain information concerning the nature of that population. Since even finite populations are usually too large to enumerate by a census, our knowledge about them must, of necessity, be derived from sampling. Experience in many fields has given ample evidence that a correctly chosen sample will be representative of the parent population, although a single sample, even if properly selected, will not duplicate *perfectly* the whole in miniature. The differences or errors, which it is virtually impossible to eliminate by any sampling technique, are known as *random sampling errors*. It has been shown that the size of these errors will be affected by the size of the sample and by the dispersion within the population of the particular characteristic in which the statistician is interested. Fortunately, these random errors tend to offset each other; their pattern of behaviour is known and this enables the probable size of the error and its expected frequency for any given sample to be determined. In consequence of such developments, sampling techniques have been evolved for numerous fields of enquiry and full confidence may be placed in the results so obtained. In the words of one authority, 'It is the development of these processes that has changed sampling from a speculative and uncertain procedure to a method having definite and determinable precision'.[1] This is only true, however, if the sample is drawn in such a way that each unit in the population has an equal or known chance of selection; *i.e.*, the sample must be random.

The Sampling Frame

This definition of a 'random' sample implies that the process of sampling is always undertaken from an accurate list or aggregate, *e.g.*, of numbered counters, of the units comprising the statistical population. This is an easy assumption to make

[1] Sampling Methods for Censuses and Surveys. F. Yates, C. Griffin.

in theory. In practice, however, it is a condition which unless fulfilled will introduce bias into the sample, since if some of the population *units* (as they are termed) are missing from the list, then clearly no sampling process can offset the fact that all units have not had an equal chance of selection in the sample. In other words, the sample is not random, it is biased. It is, therefore, of the utmost importance in sampling to ensure that the population has been clearly and unambiguously defined and then that a complete list of all units in that population is available. For example, in the survey of savings carried out in 1952 by the Institute of Statistics at Oxford in collaboration with the Social Survey the sampling 'unit' had to be defined. It was finally decided to regard a group of people who could be expected to pool their incomes and their assets and who could agree on the use of them as the 'income unit.' This was the sampling unit. Clearly, there was no ready made list of 'income units'. Even the 'households' enumerated by the Census of Population authorities would not serve, since the definition of a household in this case includes a lodger or paying guest who eats with the family unit. Such a definition would clearly not coincide with that of the 'income unit'. In the event, the survey organisers decided to draw a sample of households by selecting private addresses from the local authorities' rating lists. Special measures were needed, of course, to select one household or 'income unit' where more than one household or family shared the same premises.

The list, or it may be a card index, or even a map, of sampling units, is usually referred to as the sampling *frame*. In practical statistical work, frames are often defective and an important part of the statistician's work is to overcome the bias that may be introduced into the results as a result. The defectiveness of the frame may arise because it has been inaccurately compiled. The 1951 census of distribution enumerated all retail and wholesale selling outlets and service establishments. It is intended to use this frame for periodic sample surveys of the distributive trades. If, therefore, some shops were omitted in the census, then the sampling frame is inaccurate or incomplete. Also, it may be out of date if no method is available to record in the list any new shops which open or those which close between the census and sample survey dates. Inaccuracy in a frame arises since it is impossible in a census to ensure that all the informants are

truthful or accurate in their replies. For example, in the Ministry of Food lists based upon the ration books which were used as the sampling frame in the 1946 Family Census, some women were described as married even though in fact they were single. According to Dr Yates, 'all frames are likely to suffer to a greater or less extent from various defects,'[1] *i.e.*, inaccuracy, incompleteness, duplication and being out of date. At the outset of any enquiry it is of the utmost importance to examine the structure of the sample frame which is to be used for obtaining the sample units. Having assessed, as far as is practicable, its defects and deficiencies, methods must then be devised for overcoming them. In the Family Census each woman selected was checked to ensure that she was in fact married. The student should note that the term 'Census' is rather misleading in this case; the enquiry was actually a very large sample survey.

Random Sampling

Dr Yates has described experiments designed to demonstrate that human selection of a sample by what are known as 'haphazard' methods will not yield a random sample. By definition a random sample is one in which 'every member of the parent population has had an equal chance of being included.' The average individual is invariably biased in selection and is doubtless unconscious of the fact. To eliminate the human error in sampling alternative and more reliable methods of sample selection have been devised.

The simplest method is to number all the members of the population and then place the same number of tickets or counters in a drum and withdraw them, as in a sweepstake. The members corresponding with the numbers drawn comprise the sample. Professor Kendall has instanced a case where even this method failed to give accurate results, since a particular form of counter was more slippery than others and consequently did not turn up in its correct proportion. Another method is to select from the numbered population those members whose numbers are turned up in a table of random numbers. This is not the place for a discussion of this technique; those interested will find a clear account in Yule and Kendall.[2]

[1] F. Yates, *op. cit*

[2] An Introduction to the Theory of Statistics. Yule and Kendall. C. Griffin, pp. 376-9.

Systematic Sampling

In practice, neither of these methods is feasible unless the population is fairly small, *e.g.*, in experimental work, when a sample of a group of results is to be chosen. For most practical work it is easier to select every *n*th item in a list of the population. This method is termed *systematic*, or quasi-random sampling. Thus, if the lists comprise a population of, say, 25,000 and the sample required is 500, then the selection of every fiftieth item will yield the required sample. The starting point is determined by selecting at random a number between 1 and 50. Thus, if 37 turns up, the thirty-seventh item in the list is the first, the eighty-seventh and one-hundred-and-thirty-seventh, the second and third in the sample, and so on. Thus, when the Social Survey carried out its enquiry among ex-miners suffering from pneumokoniosis, the first name was selected by picking a number between 1 and 7 and thenceafter every seventh card was selected from the files of the pension authority. If the 'list' is in the form of a card index, there is no need to count the intervening cards if their number is large. By setting a ruler across the file, the card coinciding with a pre-determined interval, *e.g.*, every $3\frac{1}{2}$ inches, is selected.

One of the most recent official uses of systematic sampling was in 1946 by the Statistics Committee of the Royal Commission on Population. Great care was taken over the preparation of the sample comprising over one and a half million married women. By normal standards this was a fantastically large sample; most national surveys are based on about 2,500 units. The need for so large a sample was explained by the detailed breakdown of the sample which was to be made and in some 'cells' or 'boxes', as those small sub-groups are called, there would only be a few hundred units, – quite small samples. The selection was made by extracting every tenth card from the Ministry of Food's records and rejecting all males and single women. The Ministry of Food records based on the coloured cards from ration books provided a classification of the population by sex and three age groups. Women could be classified from these records as married or single. Checks were made by contacting all those selected to ensure that they were actually married, and the questionnaire was then distributed to them.

Strictly speaking, systematic or quasi-random sampling is not

truly random. This is because once the initial starting point has been determined, it follows that the remainder of the items selected for the sample are pre-determined by the constant interval. Thus, if we are selecting every twentieth address from a street list, the first is admittedly chosen by random methods, but the remainder are thenceafter pre-selected. Nevertheless, this form of sampling approximates sufficiently closely to pure random sampling to justify its widespread employment. The list or sampling frame should be checked to see whether it has been previously arranged in such a way that a particular type of unit may occur at the appropriate interval and therefore be over-represented in the sample. Generally speaking, street-lists and alphabetical lists of names are free from such bias, *i.e.* non-randomness in the arrangement of the characteristic.

These methods are based on the assumption that complete lists of the 'population' are available. For sampling the human population of this country there used to be three lists: the national register, the local authority rating lists, and the electoral roll. The first was the only complete list, but the second is very useful for sampling households. The Government Social Survey was for many years favoured by enjoying access to the National Register; all other survey organisations used the Register of Electors. When, however, the National Register was discontinued in 1952, the Social Survey was also compelled to adopt the Register of Electors as its sampling frame. It continued to use the Local Authority rating lists for such surveys as required them. Prior to this, however, the Social Survey had undertaken an enquiry into the value of the Register of Electors as a sampling frame[1] and the change was made without the difficulties which would have arisen without such knowledge of the Register's value. The following account is based upon that report.

The Electoral Register is estimated to include some 96% of the resident civilian population of England and Wales at twenty-one years and over. Since the Register is used in both parliamentary and local government elections it is possible for sampling purposes to distinguish between parliamentary constituencies and local government wards. The former are broken down into polling districts which constitute the smallest sampling unit from this frame. Although the Register is revised each

[1] The Register of Electors as a Sampling Frame, by P. J. Gray, T. Corlett, and Pamela Frankland, C.O.I., November 1950.

year so that it is reasonably up to date, any given Register is already four months old by the time it is published. It appears in March, the lists being based upon the electors' residence in the preceding November. It is effectively sixteen months old, of course, by the time the new edition appears. In other words, the Register is not continuously revised, but merely at yearly intervals, *i.e.* the November census of electors. Herein lies its main defect. Its other defect lies in the fact that a proportion of the population entitled to inclusion in the Register has not in fact been enumerated.

Unfortunately there are no accurate data available to measure the size of the error in the sampling frame arising from these omissions and unrecorded changes. The Social Survey report, however, contains an account of an investigation into this problem. It appears that about 4% of the loss is due to non-registration, while there is a further loss of ½% monthly arising from removal. Thus, whereas at the date of publication, 94% of the eligible population are included, 12 months after, *i.e.* immediately before the new Register is due to appear, only 87% of the eligible population is correctly registered. If this shortfall in numbers were evenly distributed throughout the population by reference to sex, class and income, the sampling frame would not be seriously biassed. The authors of this report estimate that the 4% initial loss, *i.e.* due to non-registration is not merely relatively small in relation to the whole, but is probably unbiassed in so far as it is spread over all groups of the population. The ½% monthly loss by removal is more serious, since it appears that a high proportion of the removals are accounted for by the under-thirties, so that the age distribution of the population remaining within the sampling frame is slightly distorted. The report concludes that the current Register of Electors 'can be used with confidence as a sampling frame if some procedure to deal with "moves" can be evolved'.

Stratified Sampling

So far it has been assumed that the population to be sampled consists of a single homogeneous group, *e.g.*, ex-miners disabled by pneumokoniosis. In many surveys the population is far from homogeneous, but markedly heterogeneous. This applies to the adult population of a country, which comprises men and

women in different age groups, in different social circumstances, and so on. Because of these differences in background, individual members of the population being surveyed to assess opinion about say road accidents, will have divergent views on the problem. As was pointed out above, all the individual members of the population may be regarded as of that population. They all share the common characteristic of holding views on the problem, but from experience it is known that certain social groups into which the population may be logically divided will think differently from other groups. If the population lists are classified into groups suitable for this particular survey, clearly a better reflection of these views will be obtained by sampling from each group, in proportion to the size of that group in the whole population. Each group will then be represented in the correct proportion within the sample. Such a sample is known as a *stratified* sample; in short, we speak of population *strata* and not groups.

It was argued earlier that even if the population is not stratified, any random sample will reproduce the distribution of the characteristic within the population. Stratification of the sample is derived automatically. Stratified samples can be drawn without first stratifying the population list and selecting from each stratum. Provided the relative sizes of the strata one to another are known, the sample members can be divided among the strata as they are drawn. As soon as the quota for any stratum is complete, any further items of that type are rejected and sampling continues until each stratum has its quota. This method will probably entail sampling a larger number than would be necessary if the population had been classified into its various strata at the outset. If the population can be so classified, then a stratified sample, *i.e.*, one made up of random samples from each of the 'strata', is likely to be more representative of the population than any other sample of that size.

A little thought will reveal why a sample drawn from a previously stratified population is more likely to be 'representative' than a similar-sized random sample drawn without prior stratification of the population. When the population is stratified the statistician is in effect drawing a random sample from each stratum or homogeneous sub-population. Within each stratum

random sampling errors must be taken into account. But the composition of the total sample, as far as its distribution between the various strata is concerned, corresponds with that of the population – because the statistician has arranged it so. In the case of a simple random sample from that population, two sets of sampling errors must be taken into account. The first are those *within* each stratum – as in the case of the sub-populations or strata within a stratified sample. Further sampling errors arise in the random sample because the distribution of units as *between* the various strata in that sample may not correspond with that of the population. It is this risk which prior stratification eliminates. The simple random sample *may* yield the correct composition of units from the various strata; but we cannot be certain and therefore when the sampling error for such a sample is computed, it is always greater than in the case of an equal sized stratified sample.

In recent years the Social Survey has developed two indices which are employed for purposes of stratifying populations by reference to their socio-economic status. The first of these is the now well-known 'J' Index. The 'J', or *Juror* Index is based on the proportion of the population which possesses a jury qualification. An examination of the Electoral Register for each polling district will reveal that certain names are preceded by the letter 'J', which implies that those individuals are liable for jury service. This qualification is dependent upon ownership or occupation of property above a certain rateable value. In other words, the larger the proportion of 'J' names in an electoral district, the higher is the number of occupiers and owners of property of a rateable value over certain limits. Since such ownership or occupation is correlated with income and social status, a high value for the 'J' Index in any area reflects a corresponding social class.[1] The second is the *Industrialisation* index, which serves much the same purpose. In this case the rateable value of an area attributable to industrial hereditaments and transport undertakings is expressed as a proportion of the total rateable value of that area. The Social Survey has ascertained that in the provinces there is a significant degree of correlation between

[1] The Social Survey report on this index is contained in a paper by Gray, Corlett and Jones entitled 'The Proportion of Jurors as an Index of the Economic Status of a District', and published by the Central Office of Information. The jury qualification is obtained either by being a householder i.e. occupier of property of £30 R.V. or over in London and Middlesex and £20 elsewhere, or, by ownership of property of £10 R.V. or more anywhere.

the degree of 'industrialisation' and the proportion of the population in the highest income group. This index, however, is not suitable for the London area and an amended form of the index is therefore employed. Here the rating areas are classified according to the rateable value per head of the inhabitants within each area. The higher the *per capita* rateable value, the wealthier is the district. These indices of stratification are nowadays regularly employed by the Social Survey and other organisations in the preparation of their samples. The difficulty with stratified sampling is that it is not usual for the population lists to be stratified. In every survey, therefore, the sampling units may have to be stratified in accordance with that particular factor which is relevant to that survey. Reference has already been made to the 'J' and the 'industrialisation' indices. Stratification may be based on rateable value per head, on the population per square mile in some sparsely populated areas, or by reference to size, *e.g.* population of districts. The stratification factor will depend on what type of stratification will be most useful for the particular survey. Professor Kendall stratified the sample used in the Readership Survey carried out for the Institute of Practitioners in Advertising by reference to the ratio of Labour to non-Labour votes in parliamentary constituencies. The higher the proportion of the Labour vote, the lower the social class of that constituency.

The advantages of prior stratification of, for example, the polling wards in parliamentary constituencies will be apparent. By arranging the wards in some order, *e.g.* by reference to the 'J' index, the required number of wards from which the sampling units will be drawn at random can be so selected that each type of ward is represented. In other words, the sample of wards will include wealthy, middle-class, and working-class wards, assuming all three types are in that particular population of wards. Pure random sampling, as was pointed out above, *may* produce a sample containing only one type of ward.

Multi-Stage Sampling

An especially useful method of sampling is known as *multi-stage* sampling, *i.e.*, the sample is prepared by stages. The population is divided into a number of large sampling units, each of which in turn is divided into smaller units, and so on. A random sample is taken of the large units at the first stage and from those

selected a further random sample, *i.e.* the second stage, is collected of the smaller units.

Although population lists of voters and households are available, the selection of a sample by quasi-random sampling would be a lengthy task. This, however, is a minor consideration compared with the time and expense which would be entailed in contacting the sample units, *i.e.*, individuals, selected in such a way. In all likelihood, the interviewers would find themselves seeking out their respondents all over the country from Lands End to John o' Groats, with a mere handful of interviews at most in any single county borough. In other words economic considerations completely outweigh any advantages that simple (single stage) random sampling may possess. Common sense suggests that some grouping or concentration of interviews in the various parts of the country should be possible, without destroying the randomness of the sample, that is its representativeness. For example, instead of say 60 interviews being dispersed all over the West Riding as could be the case if random sampling had been used, would anything be lost if those interviews were concentrated in say two towns in that area, especially if the towns in question were selected from all the towns in the area by random methods. After all, are the inhabitants of Halifax and Huddersfield so different from those in Batley, Dewsbury, etc., that the sample would be biased if the foregoing compromise were to be adopted? This seems highly improbable and the economic advantages of this compromise solution, which is the basis of multi-stage sampling, make it very attractive. In other words, statistical precision is still attainable at greatly reduced cost.

The technique of multi-stage sampling can best be illustrated by following through the various stages of drawing up a sample of the adult population in England and Wales. It is customary to distribute the interviews throughout the entire country, and for this purpose the total sample is apportioned on a regional basis. The Registrar General prepares annual population estimates for *standard regions*, which are set out in Table 32 below. The actual population in each region is given, together with its proportion of the total population in England and Wales. The final column gives the number of interviews to be carried out in each region. This is derived by apportioning the total hypothetical sample of

4,000 according to the percentage figures in the preceding column. It will be realised that such a distribution of the sampling units is actually a form of stratification, in this case on a regional basis.

TABLE 32

ESTIMATED DISTRIBUTION OF POPULATION OF ENGLAND AND WALES (30TH JUNE 1954)

Standard Region	Population 000's	Percentage of Population living in Region	No. of Interviews allotted to Region
Northern	3,151	7·1	284
East and West Ridings	4,098	9·2	368
North-Western ..	6,441	14·5	580
North Midland ..	3,437	7·8	312
Midland	4,490	10·1	404
Wales	2,601	5·9	236
South-Western ..	3,065	6·9	276
Eastern	3,258	7·4	296
London*	8,319	18·8	752
South-East* ..	2,641	6·0	240
Southern . ..	2,773	6·3	252
England and Wales ..	44,274	100·0	4,000

* The Standard Region is defined as London and South Eastern, but the percentages are shown separately since the area comprises so large a part of the total population.

The next step is to apportion the interviews within each region between the urban and rural population. Here again, the Registrar General produces figures giving the proportions of each class in each region.[1] Generally speaking, in relation to the urban population the proportion of rural inhabitants is small, but from the point of view of sampling just as important as the urban quota. This process too, is really a form of stratification. Since the rural sample is usually small it is customary to concentrate such interviews in a couple of rural areas. Thus if in the North West region 10 per cent of the interviews are to take place in rural districts, then 58 out of the 580 interviews allotted to the region will be concentrated into two rural districts selected at random from the many such as Clitheroe and Ulverston. The remainder of the region's interviews will take place in urban areas.

This is done by selecting certain towns and carrying out a given number of interviews within their boundaries. There are

[1] See The Registrar General's Estimates of the Population of England and Wales: Populations of each administrative area at 30th June for any year.

two problems here. First, how many and which towns are to be selected? After all, the whole point of multi-stage sampling is to concentrate the interviews within certain towns rather than dispersing them all over the region. Second, how many interviews shall be allotted to each town? Where there are large cities with populations of well over half-a-million, then it is simplest to allot the city the number of interviews proportionate to its population. Thus, if Liverpool has approximately three-quarters of a million inhabitants, then they form about 13 per cent of all the urban population in the North West region. The *urban* quota of interviews is 580 — 58 = 522, so that Liverpool gets 13 per cent of 522, or 68 interviews. Manchester's quota can be determined in the same way. But how does one apportion the interviews between the smaller towns, of which there are usually quite a number? The principle is as follows: These urban communities are classified in order of size, *e.g.*, all towns with populations between 150–500,000; then those from 75–150,000, then 50–75,000, down to the smallest town. The aggregate population in *each group* of given-sized urban districts is then expressed as a proportion of the entire urban population in the region and this proportion applied to the number of urban interviews in the region's quota. Thus, if the towns with populations between 75–150,000 contain 11 per cent of the population, then 11 per cent, or 57 of the region's urban interviews will take place in such towns. When, however, there is a number of towns of this size it is pointless to spread the interviews over them all. This would defeat the whole purpose of multi-stage sampling. Since experience has shown that the most efficient distribution of interviewers is achieved by allotting each one about 30 interviews, then two interviewers in two towns can do the entire quota for these towns. The two towns are chosen at random from their particular group. The same procedure is adopted for each group of towns, always ensuring that interviewers are given a block of about 30 interviews in a single locality.[1]

Having decided upon the towns in which the interviews are to take place, the next stage is to determine the actual people to be

[1] In the 1953 National Food Survey and the 1957 I.P.A. readership survey the first stage sampling units were parliamentary constituencies in the twelve standard regions, the actual number of constituencies selected for each region depending on the proportion of the population in the region. The advantage of using parliamentary constituencies as first stage sampling units instead of the administrative areas such as county boroughs, non-county boroughs etc. is that they are more uniform in size and the grouping (by reference to size) of the administrative areas described above was unnecessary, although the largest towns were taken out separately as described above.

interviewed. It would be possible to select their names by quasi-random methods from the Electoral Register of the selected town, or households (if that is the sampling unit) from the local authority's rating list. The former sampling frame is more usually employed, but to ensure the maximum likelihood of a representative sample the electoral wards, which form the smallest unit within the Electoral Register, may be classified by reference to the J index. If the local authority lists are being used, then the stratification will usually be based on the industralisation index outside London and on the rateable value index within London. A given number of wards or rating areas are then selected at random and finally the names of the persons to be interviewed are selected by systematic sampling from the electoral or rating lists of those wards or areas.

It may be helpful to recapitulate briefly the preparation of such a sample which is typical of most national surveys. Note, first of all, that in this particular sample three types of stratification were introduced. Strictly speaking, they are not essential in the construction of a multi-stage sample. It would be quite feasible to sample at random from the first, second and third stage sampling units, as they are called, without stratification. However, for the reasons already given, the sampling units at each stage were stratified; initially by reference to the regional distribution of the population and in accordance with the urban/rural distribution, then according to the size of the urban areas, and lastly by reference to some index reflecting social class, *e.g.*, the J index. The actual stages in the sample were, first, the selection of the towns (and rural areas) in which the interviews were to be held. Having decided on the towns, the next stage was to sample from each of these towns the required number of wards, and finally at the third stage to select within the selected wards the individual names and addresses.

The selection of the towns or urban areas, apart from the large cities which, by virtue of their size justify their own quota of interviews, is more complex than may appear at first sight. Within each stratum or group the towns may vary quite substantially. Since the smaller towns are usually more numerous than the larger, if simple random selection were used the chances were that within any such group the smaller towns would have a greater chance of inclusion than the larger towns. The larger the

strata employed, the greater does this risk become. To overcome the danger that smaller towns will be disproportionately represented in the sample simply because they are more numerous, although in the aggregate they comprise a smaller proportion of the sampling units, a method of sampling with 'probability (of selection) proportionate to size' is used. Within each group or stratum the towns are ranged in some order, *e.g.*, of size, and their populations added cumulatively in the same way as if we had to calculate the median of such a group. Then two (if this is the number of towns to be selected in that group) numbers are selected from a table of Random Numbers and the towns within whose range those numbers fall, are selected. It will be realised that the larger towns in each group will have a larger range of numbers in the cumulative total for the group and the likelihood of the random number selected falling within their particular ranges is greater than the chance the random number would fall into the smaller range of numbers absorbed by the smaller units. In this way, the probability of selection is proportionate to size, *i.e.*, a town twice as large as another has twice as many chances of selection, just as the holder of 100 Premium Bonds has ten times the chances of winning a prize in any month than the holder of only 10 Bonds.

The weakness of multi-stage sampling such as has been outlined above, will be fairly obvious. Since random sampling errors cannot be avoided, they must accumulate at every stage, and the sampling error will be larger in a multi-stage sample than for a sample of the same size selected by single-stage stratified sampling. Since most of the present-day nation-wide surveys are based on samples collected in this manner, the problem deserves serious consideration. Although the error is larger than for single-stage stratified sampling, with widely scattered populations the cost per completed questionnaire is sufficiently lower to offset the disadvantage cited. This weakness can be partially eliminated by increasing the size of the sample to obtain the same degree of accuracy as would be obtained by a smaller stratified sample. The advantages of multi-stage sampling are extremely important. For nation-wide surveys it is the only method which is administratively practicable, especially as it only becomes necessary at the final stage to use the lists of the population to be sampled. In the first and second stages

sampling from national population lists is unnecessary, since it is the administrative counties and then cities and towns which are sampled. Only for those *selected* areas are sampling frames necessary.

Quota Sampling

To economise in time and cost, American practice has provided a technique known as *quota* sampling. In a nation-wide survey the size and composition of the sample and the stratification are all determined at the centre. The interviewer, instead of receiving a list of names and addresses to be interviewed, receives a list of individuals classified by social types, *e.g.*, one professional man over 45, two labourers, two clerical workers one male and one female both under 30 years, and so on. The interviewer can then select at her own discretion all the required interviews without considerable travelling and possible time-wasting. The value of knowing the strata comprising the population can now be appreciated since quota sampling would be impossible without this knowledge. The classification employed is usually quite simple; for example, the stratification employed by the Hulton Readership Survey shown below. The student may care to compare this particular classification with that used by the National Food Survey. The latter is given in detail in the report on Domestic Food Consumption and Expenditure in 1950, published by H.M. Stationery Office in 1952.

This classification probably appears deceptively simple and rather arbitrary. It is, furthermore, based on rather doubtful statistical data. The Registrar General does classify the population by reference both to social class and socio-economic status, but the divisions are broad and the data can at best serve as a rough guide to the preparation of a stratified quota sample. In practice the interviewer is provided with detailed schedules of the criteria by which the informant's class was to be judged, *e.g.*, occupation, possession of a car, house and so on. Each interviewer is given the 'quota' of interviews which he is to obtain with so many respondents of either sex from these various classes. This, in a sense, is no more than the prior stratification of the population; but whereas in all the random samples described above, the interviewer is given specific individuals to

The Social Classes*

(*Source: The Hulton Readership Survey* 1954)

Class	Description	Brief Description	Percentage of Families in each Class	Usual Income Level of Head of Household
A	The Well-to-do	Doctor, chartered accountant, barrister, solicitor, civil servant in the administrative grades, town clerk, headmaster of a public school, university professor, hospital matron, county welfare officer (woman), stockbroker, owner (or senior executive) of a large business, manager of a large branch of a bank, farmer of a large farm.	4	Over £1,300 p.a.
B	The Middle Class	Surveyor, estate agent, qualified engineer, industrial scientist, veterinary surgeon, chief librarian, headmaster of a smaller school, vicar, civil servant in the senior executive grades, other bank managers, owner (or senior executive) of a medium sized business (factory or shop), farmer of a medium sized farm.	8	£800 to £1,300 p.a.
C	The Lower Middle Class	Civil servants in the intermediate executive grade, bank clerk, managing clerk, shorthand typist, teacher, library assistant, physiotherapist, nursing sister, commercial traveller, foreman, station master, owner (or manager) of a small business or shop, farmer of a small farm.	17	£450 to £800 p.a.
D	The Working Class	Skilled and semi-skilled manual workers: bricklayer, painter, plumber, fitter, machine operator, farm worker, garage mechanic, metal worker, foundryman, toolsetter, press worker, riveter, miner, postman, policeman, bus-driver, railway worker, shop assistant, junior clerk-typist, barman, waiter.	64	£250 to £650 p.a.
E	The Poor	Unskilled manual workers: labourer, cleaner, market porter, charwoman, packer, domestic help, lampman, bargeman, snack-bar attendant, scrap sorter. Also old age pensioners who are almost solely dependentupon the pension as a source of income.	7	Under £250 p.a.

* In the preceding table emphasis has been placed on the 'social' rather than on the 'economic' standard of classification, since information about income is difficult to obtain. The assessment of the informants has been made by interviewers on the basis of their appearance, speech, occupation, type of house and district in which they live (in the case of persons interviewed at home); in fact, all those subjective characteristics by which we are accustomed to make a mental assessment of people's social and economic position.

interview, in quota sampling the choice of the actual sample-units, *i.e.*, respondents, is left to the interviewer.

The weaknesses of this technique are all too apparent. It has been established that truly random selection by any individual without some pre-arranged procedure is impossible. Interviewer bias in the absence of direct control in selection of their respondents is unavoidable. And since the sample is not truly random, the significance tests for reliability of the results cannot legitimately be applied. This is a vital weakness, since the whole essence of random sampling is that it enables the statistician to state the probable accuracy of his results within clearly defined limits. To the extent that much of the interviewing undertaken for the commercial organisations may take place outside the home, bias may be introduced into the result by an inadequate representation of those people who spend most of their time at home, for example, the very old, or housewives with large families. To overcome this danger the interviewer is often given a quota in which he is given a clear indication of where he will find suitable respondents. He may be asked to contact so many shop-workers, or transport workers, farm workers, office workers and housewives. In each case the adjective indicates the most suitable place in which the respondent may be contacted. This more detailed classification, as Mr F. Edwards of the British Market Research Bureau has pointed out in his paper on sampling methods,[1] is particularly important with quota sampling since using the simplest form of socio-economic classification approximately 70% of the population comes within the group known as 'working class'. It is imperative therefore that this large group should be broken down to enable the interviewer to collect the appropriate respondents thereby ensuring that the sample is representative of the entire population, or that part of it known as the 'working class'. In brief, the degree and precision with which the 'quota' sample can be stratified forms the key to the reliability of the results.

[1] Sampling Methods, Part II in 'Modern Sample Survey Methods' published by the Association of Incorporated Statisticians.

It is noteworthy that the Government Social Survey only employed quota sampling for some of its enquiries in the first two years of its existence, but thereafter dispensed with the method for the reasons given above. But those employing quota sampling, *i.e.*, most commercial market research agencies and public opinion polls, argue that it is adequate for their purpose. Against the above criticisms those employing quota sampling argue that their interviewers are properly trained, carefully briefed and keenly aware of the problems involved. This is certainly true of one well-known international business organisation in this country, which has a large market research organisation employing full-time interviewers. Even the Social Survey only employs part-time interviewers, although great pains are taken to train and then test them. Furthermore, the results from every interviewer are checked and recorded over a long period and any bias noted. Finally, the main argument, since given adequate resources both types of sampling could be carried out with fully trained staff, is that the commercial organisations do not require the same degree of accuracy as the Government survey unit on whose results administrative action affecting the whole population may be based.

Even if the bias arising from quota sampling (assuming it could be measured) is greater than that in random sampling, it need not affect the final results of the survey very greatly. The reason for this statement is to be found in the fact that by far and away the major errors in any survey are committed at the interviewing stage where, unfortunately, they frequently remain undetected. Relatively to this type of error, the bias introduced by a quota sample may be small. Furthermore, it should be noted that due to non-response and failure to contact selected respondents, even a random sample may suffer from bias. This is because the 'non-respondents' may form a particular group and may not be the same 'type' as those respondents who were interviewed. This point is very important and is discussed further on p.230. For all practical purposes, therefore, the weaknesses of quota sampling are not as serious as they appear at first sight.

Much of the criticism levelled against quota sampling is in part a relic of the early days of American experience before interviewers were trained and all the difficulties of the sampling technique appreciated. The tightening up of the organisation of

commercial agencies, the institution of checks on interviewers' work, and an analysis at intervals of their results, has unquestionably overcome the *worst* dangers. For example, in one organisation using quota sampling it has been stated that the supervisors call back on 10% of the respondents. Another organisation makes a postal check on 10% of the interviews conducted by their interviewers, and carries out an occasional 100% check on individual interviewers. As the author of a paper on quota sampling has pointed out, such checks in themselves may not be particularly effective, but the psychological effect of their existence upon interviewers is undoubtedly considerable.[1] In the meantime, public opinion polls employing quota sampling can at least point to their record and show that their method has usually provided accurate results, despite the adverse publicity given to one or two failures.

Cluster Sampling

As we have seen, simple random sampling does not yield such a high degree of precision as stratified sampling, while multistage sampling is almost as much the product of economic as statistical considerations. But in all these samples, what is termed the *primary sampling unit* is the individual unit, *e.g.*, an adult or a household. The larger the number of primary sampling units, the larger will be the cost of a survey. Sometimes the cost factor may necessitate a different form of sampling whereby interviewers concentrate all their interviews in a relatively small number of areas or groups. Suppose that a survey is being carried out over a large area in which the population is extremely dispersed. A random sample would be quite impracticable. Alternatively, the survey may be concerned with measuring the number of homes with refrigerators in a large area for which there exists no list of these homes, *i.e.*, there is no sampling frame. To carry out a census to derive a sampling frame would be very expensive indeed. This was in fact done in the United Kingdom in 1950 as a preliminary to the Census of Distribution which was later to provide a sampling frame for future periodic sample surveys. But usually, where no list or frame exists, random sampling is impossible. Furthermore, where the sampling units are widely scattered, the costs of a pure random

[1] Quota Sampling, C.A. Moser. J.R.S.S. Part III, 1952. See also 'An Experimental Study of Quota Sampling,' by C. A. Moser and A. Stuart J.R.S.S. Part IV, 1953.

sample could be considerable. If, however, a few blocks of dwelling houses or localities were selected at random and *every* individual in each block interviewed, then if the blocks when put together form a sample which constitutes a representative group of the population, the statistician will have achieved his objective, *i.e.*, a random sample of the entire population.

To meet the problem of costs or inadequate sampling frames in the United States this method of *cluster* sampling, sometimes known as *area* sampling, has been devised. By the use of map references, the entire area to be surveyed is broken down into smaller areas and a few of these areas are selected by random methods. The primary sampling unit is then no longer the individual but is a group of individuals or households to be found within the slected area. Such groups are termed 'clusters'. Within each area selected, sometimes every unit, *e.g.*, household, is interviewed. Sometimes it is only a proportion, say, one in four households. Nothing need be known in advance about the area, the number or type of sampling units in it, but by following these procedures the chances of inclusion in the sample can be made the same for all individual units to be found within the area.

The basic problem with this type of sample is whether or not the units within the clusters are homogeneous. The danger undoubtedly exists that clusters often tend to comprise people with similar characteristics and since the statistician is picking out only a few clusters, he may find himself with a biased or non-representative sample. If the individual clusters are heterogeneous, *i.e.*, made up of all types of individual, then the final collection of 'clusters' may well constitute a random sample. If, however, the clusters are highly homogeneous in their composition, the reverse is true. In other words, whereas the statistician who wants to be able to stratify his sample is concerned to ensure that each stratum in the population is homogeneous, the same statistician using cluster sampling would prefer the areas, *i.e.*, strata from which he is sampling to be heterogeneous. In practice, the statistician using cluster sampling is well advised to take a sample consisting of a large number of small clusters rather than a similar sample containing only a few large clusters.

Cluster sampling has been evolved in the United States because it permits surveys to be undertaken with low costs and

also because adequate sampling frames for the relevant populations are not always available. In this country we are not confronted with the problems of widely dispersed populations. In the United Kingdom, area sampling was used in the Census of Woodlands in 1942 for which it was eminently suited.

The Sample Size

The point has already been made that the costs of a sample survey are directly related to the size of the sample used. The object of sample design, as it is called, is to maximise the degree of accuracy or precision in the sample results for any given outlay. We have seen that a stratified sample will give a greater degree of precision than a simple random sample; while both multi-stage and, to an even greater extent, cluster sampling represent compromises between statistical and economic considerations. Whichever type of sample design is used in a survey, the inevitable question arises as to the size of sample to be taken. If one asks the simple question, 'what is the appropriate sized sample for a particular survey', the answer is invariably, 'the largest practicable.' We have seen that every increase in the sample size brings with it some increase in the precision of the sample estimate. The point has also been made, but it bears repetition since it often puzzles the layman, that the size of the population from which the sample is to be drawn is quite irrelevant.

The key to the question as to the appropriate size of a sample is determined by the results required. Let us assume that the leaders of a political party want to know the proportion of the electorate which approves their particular policy. The statistician may inform them after an opinion poll has been taken, that he is 95 per cent certain that between 40 and 50 per cent of the electorate support the party. This is clearly of little value; 50 per cent means victory at the polls, the figure of 40 means defeat. In the example quoted, it is quite clear that the standard error of the percentage is $2\frac{1}{2}$ per cent since the 95 per cent level of confidence sets limits of twice that error about the sample statistic of 45 per cent. To give more precise results at the same level of confidence, the statistician must take a larger sample and his clients must therefore pay more for his work. Suppose the clients will be satisfied to know with 95 per cent confidence within one

per cent either way the proportion of the electorate supporting them. In other words, the sample must yield a standard error of 0·5 per cent. From the formula for the standard error of a percentage, *s.e.* $\% = \sqrt{\frac{pq}{n}}$ we can by substitution arrive at the required sample. Thus:

$$0{\cdot}5\% = \sqrt{\frac{45 \times 55}{x}}$$

$$0{\cdot}25 = \frac{45 \times 55}{x}$$

$$0{\cdot}25x = 2{,}475$$

$$\therefore\ x = 9{,}900$$

Strictly speaking, the above formula applies only to a simple random sample. As has been explained, the gains in precision from prior stratification of the population or the sample are considerable. But the formula for deriving the standard error of a stratified sample is much more complex. It consists largely of summing the standard errors within each of the strata making up the sample. Similarly, the calculation of the standard error of a multi-stage sample is complicated by the fact that at each stage a random sample is taken of the relevant sampling units and these standard errors accumulate. Generally speaking, however, the above simple formula based upon the standard error of a proportion gives a useful and easily calculated guide to the maximum sample required in a survey which is concerned to ascertain the extent to which the population possesses a particular attribute, *e.g.*, watches I.T.V. or votes Liberal, etc. When the statistician is dealing with variables, *e.g.*, the average income of members of a given population, then a different formula is required.

The main object of the foregoing section is to impress upon the reader that sample size has nothing to do with the size of the population and that, in a sense, the statistician works backwards from his probable results to decide upon the required sample. The important consideration is the degree of precision required in the results. The importance of costs has been much stressed, but no statistician will subordinate statistical considerations to

considerations of finance. His function is to advise his clients as to the best and cheapest way of obtaining the information they require. If they are not prepared to meet the cost of what the statistician considers to be the minimum sample required to yield the information they have asked for, then he will advise them that to undertake the enquiry will merely waste their money.

Conclusions

In this chapter an attempt has been made to describe in simple terms the main types of sample which are currently employed in survey work. Reference has also been made to the considerations which may determine the actual sample design employed by the statistician. Great emphasis has been placed upon the need for random selection of the sampling units. Only if this rule is observed can the precision of sample statistics be measured by calculating their standard error. Considerations of economy and time have led to the widespread adoption by many commercial market research agencies of quota sampling. From the statistical point of view this method is inferior to random sampling; but random sampling is itself subject to other weaknesses. Provided the data are available to enable quota samples to be stratified in some detail, the results achieved are undoubtedly good enough for the purposes for which quota samples are generally used.

CHAPTER XII

REGRESSION AND CORRELATION

So far we have been discussing the description and analysis of one variable. For example, the weekly turnover of twenty retail shops can be set out in tabular form as follows and for this distribution it is possible to calculate the arithmetic mean and the standard deviation.

Retail Shop		1	2	3	4	5	6	7	8	9	10
Weekly Turnover	£	150	200	210	230	260	280	300	320	350	370
Retail Shop		11	12	13	14	15	16	17	18	19	20
Weekly Turnover	£	380	400	410	430	460	470	480	500	520	540

A similar tabulation can be constructed to show the gross profit of each of the same twenty shops and the same statistics calculated. There would, however, probably be some relationship between turnover and the amount of profit. We might ask, for example, whether the profit increases constantly with the turnover? In order to answer this question the data can be arranged, as a preliminary to further analysis, in the form of a table such as that above together with the following data relating to profits.

Retail Shop		1	2	3	4	5	6	7	8	9	10
Profit	£	30	35	40	45	50	50	60	65	70	70
Retail Shop		11	12	13	14	15	16	17	18	19	20
Profit	£	80	75	85	90	100	80	90	100	110	115

There are, however, rather too many figures to judge to what extent turnover and profit are directly connected, *i.e.*, that a given increase in sales is accompanied by a specific increase in

profits. A simple device for examining the data so as to bring out any relationship between the two variables is the so-called *scatter* diagram.

This is an ordinary graph on which turnover is measured along the base and the profit against the vertical axis. For each shop there are two values which together locate a point on the graph. The twenty pairs of values are plotted on the graph depicted in Figure 19A. It is immediately apparent that the points plotted form a clear pattern diagonally across the graph from the bottom left-hand corner to the top right. Such a pattern signifies that the amount of profit tends to rise with the sales volume of each shop.

Lines of Regression

It is possible to define the approximate relationship between profit and turnover in mathematical terms by means of an equation. Any straight line conforming to the path of the points plotted in the graph can be defined by an equation of the form $y = a + bx$. Given this equation and the values of a and b, which are termed constants, then by substituting in the equation any value of x, we can derive a value for y. Unfortunately, in this particular case, the points plotted do not lie exactly along such a line. They are scattered about that line, some above and some below.

Now rearrange the data given above in the form of a frequency distribution. For any given turnover, what is the average profit? This can be obtained by setting down the figures of turnover and profits in two separate groups:

Turnover x	Observed profit y	Average profit	Turnover x	Observed profit y	Average profit
150—	30	30	400—	75, 85, 90	83⅓
200—	35, 40, 45	40	450—	100, 80, 90	90
250—	50, 50	50	500—	100, 110	105
300—	60, 65	62½	550—	120	120
350—	70, 70, 80	73⅓			

and for each class of turnover we can derive an 'average' profit. If these average figures are now plotted with the mid-points of the corresponding classes of turnover, it will be seen from

Figure 19B that the approximation of the plotted points to a straight line is much better. The line that has been drawn on the basis of these points, *i.e.*, through the means of each group of y values corresponding to each value of x, is drawn in such a way that the squares of the vertical distances between each of the points and the line under or over the point are at a minimum. Such a line is called a *regression* line because it is derived from the equation which defines the regression of y upon x. In terms of our data, it measures the extent to which y, *i.e.*, profit, is determined by or dependent upon the value of x, *i.e.*, the turnover.

The rather curious term regression needs some explanation since it has no obvious relevance to the data. This type of regression analysis was developed by Sir Francis Galton towards the end of the 19th century and the original data he used related to the heights of fathers and sons. He found that on average tall fathers had tall sons, but the sons tended to 'regress' to the average male height. The term has remained in use for all types of data ever since.

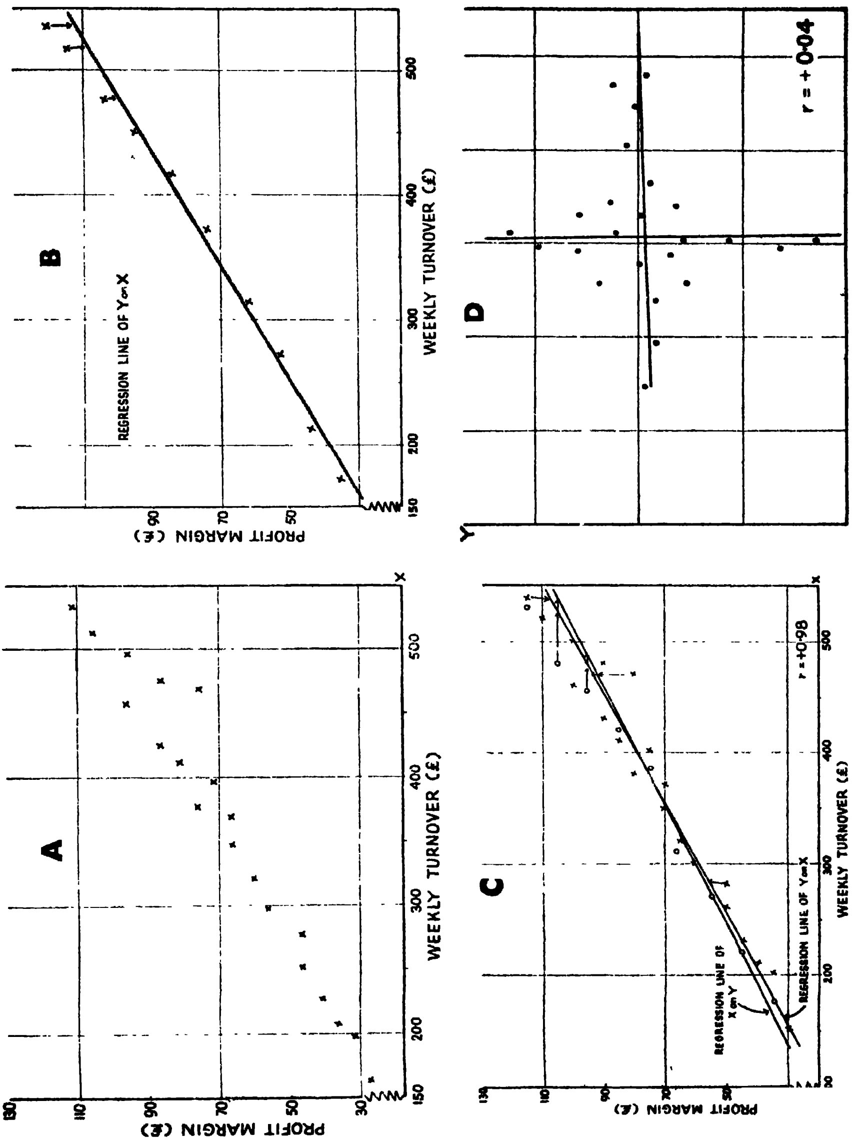
So far the data have been studied from the point of view that turnover determines the rate of profit. The relationship between such variables is not always such that the causal effect is apparent. It is, for example, reasonable to assume that profits depend upon turnover, but the causal relationship between the increase in cancer and, say, the expansion of motoring, which if plotted might reveal a marked association, is by no means obvious or certain. It is therefore sometimes practicable to reverse the two variables. It might be argued that the level of profit determines the turnover[1] and in terms of our regression equation, profit should be represented by the *independent* variable x and turnover would then be the *dependent* variable y. Then we re-classify the original data so that for given rates of profit we can calculate the average turnover. This is done in the table below and the values for x and y are plotted on the scatter diagram as small circles.

For these points as well, it is possible to fit a regression line. In this case, however, it is so drawn that the squares of the *horizontal* differences between the points and the regression line are minimised. This particular line is known as the regression

[1] For example, in those shops where profits are high every effort may be made to expand turnover and the scale of business.

Figure 19

REGRESSION LINES



Profits x	Observed Turnover y	Average Turnover	Profits x	Observed Turnover y	Average Turnover
30—	150, 200	175	80—	380, 410, 470	420
40—	210, 230	220	90—	430, 480	455
50—	260, 280	270	100—	460, 500	480
60—	300, 320	310	110—	520	520
70—	350, 370, 400	373	120—	550	550

line of x upon y. It yields similar information to the other regression line of y upon x.

It will be seen from Figure 19C that these two lines do not coincide, although they intersect each other. But suppose, when the first regression line was drawn, all the points lay along its path. In other words, the *vertical* distances between the line and the observed values of the dependent variable for each value of the independent variable were zero. If we were then to draw the other regression line so that the *horizontal* differences were at a minimum, where would it lie? Clearly, it would coincide with the first line; in other words, the regression lines, and therefore the two regression equations, are identical. But, this situation will only arise if the relationship between x and y is perfect, *i.e.*, any given movement in x is invariably accompanied, by a proportionate movement in y. It then becomes possible to predict from the regression equation the value of y which corresponds with any value that x may take. Such a perfect relationship, or degree of association as it is sometimes termed, is the exception rather than the rule. The points do not usually lie along a single line and therefore for any set of plotted points it becomes possible to draw two regression lines. At one extreme these lines coincide, that is when the relationship between the two variables is perfect. In Graph C of Figure 19 the two regression lines are very close and therefore the degree of association is high. This is given by the value of r which has a maximum value of unity and a minimum of zero.[1] At the other extreme in Graph D, where there is virtually no association whatsoever between the two variables, the regression lines are at right angles to each other and the value of r is virtually nil. A note explaining the method of calculating these regression lines is given at the end of this chapter.

[1] This is explained below.

The regression lines always cross one another at the means of the two distributions of x and y values. If the mean turnover is calculated it is found to be £363 and the mean profit is £72. Reference to the diagram will show that these values mark the point of intersection. This fact of intersection at the respective means of the two variables is very important. The degree of association between the variables is measured by the extent to which the deviations for each pair of observed x and y values (which form a point on the graph) correspond, the deviations for each pair of x and y values being measured from the respective means of the two variables. If the relationship between the deviation for a given value of x and the corresponding deviation for the related value of y is fairly regular, then it follows that the two variables are co-related, *i.e.*, movements in one variable are associated with given changes in the other. The importance of measuring the deviations of values of both variables from their respective means will be more apparent when we calculate the degree of association, or what is known as the *coefficient of correlation* represented by the letter r.

The Coefficient of Correlation

In the above illustration the relationship between x and y was such that the pairs of observed values increased together or declined together. In such a case the correlation is described as positive or *direct*. When an increase in one variable is accompanied by a fall in the other, for example, an increase in family incomes is accompanied by a fall in the consumption of bread and potatoes, then the correlation is known as *inverse* or negative. The distinction between the two types of relationship is always indicated by the sign, plus or minus, placed before the value of the coefficient of correlation. This coefficient is zero when there is no association between x and y, *i.e.*, when the regression lines would be at right angles to each other (Graph D). When the regression lines coincide, however, *i.e.*, perfect association, the coefficient has a value of unity. The formula for determining the value of the coefficient is such that the latter will always lie between zero and unity, either positive or negative. It can never be greater than 1. Therefore, if the student finds at the end of an exercise that his coefficient exceeds unity, then he will know he has made a mistake in the calculations.

As already explained, the coefficient of correlation is a measure of association. Sometimes it is a measure of inter-dependence, in which case we say the relationship is causal. It differs from the equation defining either of the lines of regression in so far as the latter define a unique relationship from which, given the change in (say) X, it is possible to compute the most probable change that will follow in Y. The coefficient of correlation merely indicates the closeness or intensity of the association without defining it. For example, the correlation coefficient of + 0·9 given by the data relating to the 20 retail shops suggests that there is a fairly close relationship between the turnover and profit. If the coefficient had been only + 0·3 we should have deduced that the relationship was not sufficiently evident – at least from the data available – to justify any use of the relationship for analysis or prediction. Generally speaking, it is customary to compute the correlation coefficient and to ignore the lines of regression and the regression equations unless the relationship between the variables is such that it may reasonably be summarised in the form of a mathematical equation. The next few sections of this chapter are devoted to explaining methods of deriving the value of r. The coefficient of correlation, like any other statistic, is derived from sample data. The larger the sample the more reliable the statistic as an estimate of the corresponding parameter. But in order to simplify the exposition and keep the calculations to a minimum the illustrations are based on typical examination questions rather than realistic data. This approach is fully justified because although references in the social science literature to the employment of correlation technique are frequent, it is in practice an extremely difficult statistic to interpret.

Calculating the Coefficient

When the relationship between two variables is linear and the data have not been grouped, and only in such cases, the correlation is computed by means of the following formula:

$$r = \frac{\Sigma\, xy}{N\sigma_x\, \sigma_y}$$

where the symbols x and y in the numerator (*i.e.*, the upper part of the fraction) represent not single values of x and y, but the

deviations of all x and y values from their respective means, in the same way as we used the deviations from the mean when calculating standard deviations by the short method (See p. 126). To remind us of the meaning of the above formula for R, we can write the numerator as $\Sigma(x - \bar{x})(y - \bar{y})$ where x is any single value of that variable and $\bar{x}$ is the mean of all the x values.[1] The same applies to y. Note that these 'deviations' will also be used for calculating the two standard deviations which the formula requires.

This formula or method is sometimes referred to as the product-moment, or the Pearsonian coefficient; the first following the method of calculation and the second the name of its discoverer, Karl Pearson.

In the following illustration the data relate to the turnover and the profit margin of the sample of shops discussed earlier. It is required to calculate the correlation between the size of the weekly turnover and profit margin. The data are given in

CALCULATION OF r BY PRODUCT-MOMENT METHOD

1 Retailer No.	2 Weekly Turnover (x) £	3 Profit Margin (y) £	4 $(x-\bar{x})$	5 $(y-\bar{y})$	6 $(x-\bar{x})^2$	7 $(y-\bar{y})^2$	8 $(x-\bar{x})(y-\bar{y})$
1	150	30	— 213	— 42	45,369	1,764	+ 8,946
2	200	35	— 163	— 37	26,569	1,369	+ 6,031
3	210	40	— 153	— 32	23,409	1,024	+ 4,896
4	230	45	— 133	— 27	17,689	729	+ 3,591
5	260	50	— 103	— 22	10,609	484	+ 2,266
6	280	50	— 83	— 22	6,889	484	+ 1,826
7	300	60	— 63	— 12	3,969	144	+ 756
8	320	65	— 43	— 7	1,849	49	+ 301
9	350	70	— 13	— 2	169	4	+ 26
10	370	70	+ 7	— 2	49	4	— 14
11	380	80	+ 17	+ 8	289	64	+ 136
12	400	75	+ 37	+ 3	1,369	9	+ 111
13	410	85	+ 47	+ 13	2,209	169	+ 611
14	430	90	+ 67	+ 18	4,489	324	+ 1,206
15	460	100	+ 97	+ 28	9,409	784	+ 2,716
16	470	80	+ 107	+ 8	11,449	64	+ 856
17	480	90	+ 117	+ 18	13,689	324	+ 2,106
18	500	100	+ 137	+ 28	18,769	784	+ 3,836
19	520	110	+ 157	+ 38	24,649	1,444	+ 5,966
20	540	115	+ 177	+ 43	31,329	1,849	+ 7,611
	7,260	1,440	0	0	254,220	11,870	+ 53,780

[1] The symbol σ means the 'sum of' all these products.

$$\bar{x} = \frac{7{,}260}{20} = 363. \quad \bar{y} = \frac{1{,}440}{20} = 72.$$

$$\sigma x = \sqrt{\frac{254{,}220}{20}} = 112{\cdot}7. \quad \sigma y = \sqrt{\frac{11{,}870}{20}} = 24{\cdot}4.$$

$$r = \frac{\Sigma(x-\bar{x})\,(y-\bar{y})}{N\sigma x\, \sigma y} = \frac{53{,}780}{20 \times 112{\cdot}7 \times 24{\cdot}4}$$

$$= \frac{53{,}780}{54{,}997{\cdot}6} = 0{\cdot}98.$$

columns 1, 2 and 3 and the calculations may best be followed if set out in stages. The true means of both distributions are easily found. In the case of the weekly turnover it is £363, and in the case of profit, £72. In columns 4 and 5 the deviations from the true means of the numbers of the items in each variable are set out. In columns 6 and 7 the deviations given in columns 4 and 5 have been squared and summed. The calculation so far is no more than would be required for deriving the standard deviations of any two distributions. Since the deviations of the individual values in both series have been measured from their true means their sum is zero (see columns 4 and 5). Column 8 provides the sum of the cross products. It will be seen that the deviations given in columns 4 and 5 are multiplied together and the products aggregated. Their sum, which forms the numerator in the formula for calculating the coefficient is often referred to as the 'co-variance'. The next stage is to substitute these values in the appropriate formula. The numerator is given by the sum of the cross-products in column 8 and is divided by the number of pairs of items, *i.e.* 20, and the product of the two standard deviations. The two standard deviations are easily derived from the data given in columns 6 and 7. It will be noted that the working in each case is from the true means: no correction is necessary. Substituting the appropriate values in the formula for the correlation coefficient, the value of the coefficient is derived by simple arithmetic.

The numerator in the formula is based on the need to provide a meaning for 'high' and 'low' values of the coefficient. All the values of each variable must be related to a comparable value; in this case the means of the two distributions are the norm for each pair of values. Provided the pairs follow some pattern in respect of their divergence from their respective means – which

will be the case if the two values are related – then positive deviations from the mean will be fairly regularly associated either with similar positive deviations of the other variable when r will also be positive, or with negative deviations, when r will be negative. It will be seen from the example below that if high and low values of X and Y are indiscriminately associated (*i.e.* the paired variations from the respective means of both X and Y values are not at least fairly consistent in direction) then the sum of the products of the deviations will be small. Therefore r too, will tend to be small.[1]

		I*a*			I*b*		
X	$(x-\bar{x})$	Y	$(y-\bar{y})$	$(x-\bar{x})(y-\bar{y})$	Y	$(y-\bar{y})$	$(x-\bar{x})(y-\bar{y})$
3	−6	7	−12	+72	28	+9	−54
5	−4	14	−5	+20	24	+5	−20
7	−2	20	+1	−2	21	+2	−4
9	0	21	+2	0	20	+1	0
13	+4	24	+5	+20	14	−5	−20
17	+8	28	+9	+72	7	−12	−96
6)54		6)114		+182	6)114		−194
AM=9		AM=19			AM=19		

		II.		
X	$(x-\bar{x})$	Y	$(y-\bar{y})$	$(x-\bar{x})(y-\bar{y})$
3	−6	20	+1	−6
5	−4	14	−5	+20
7	−2	28	+9	−18
9	0	21	+2	0
13	+4	7	−12	−48
17	+8	24	+5	+40
6)54		6)114		−12
AM=9		AM=19		

In numerical terms the sum of the cross-products will be large the larger the absolute value of the deviations. The case of the standard deviation may be recalled; this is expressed in the units of the distribution and if the members are large the S.D. is large, *e.g.* the S.D. of the weights of adult males in absolute terms is

[1] Note that the values of X are unchanged for all three sets of values of Y. In Ia, the high values of Y are related to high values of X; in Ib the reverse. In both cases the value $\sum xy$ is large in arithmetic or absolute terms. In case II there is no pattern about the values of Y in relation to X and $\sum xy$ is a negligible quantity, therefore r would also be very small. The standard deviations of the 3 distributions of Y given above will, of course, be the same in each case.

greater than the S.D. of the weights of adult females; but if related to their respective means *e.g.* as with the coefficient of variation, the dispersion of weight among females is greater. For similar reasons, in the product-moment formula, the sum of the cross-products are related to the standard deviations of the two distributions. In effect, the cross-products are 'standardised' so that the value of r is no longer influenced by the absolute value of the factors in the cross-products; note too, that in consequence the value of r is a coefficient that has no dimensions.

Students often complain that they cannot remember the formula for calculating the coefficient of correlation. This is understandable, but if they remember the basis of the statistic it should be possible to work it out. It has been explained that the value of r is dependent on the degree to which deviations of x and y move in sympathy. As was shown on p.209 if they do not correspond, then the sum of the cross-products is small; if they do, then the sum of the cross-products is high.

On the other hand, the sum of the cross-products may be high because the actual values, and therefore the deviations, are large in absolute terms, *e.g.*, they may be expressed in pounds instead of ounces. Secondly, their sum will tend to be greater, the larger the number of pairs of variables there are. Obviously a statistic which is affected both by the number of observations and the absolute size of the variates is unsatisfactory. This weakness is overcome by relating the sum of the cross-products (sometimes called the 'co-variance') to (*a*) the number of paired observations, *i.e.*, N and (*b*) a measure of the actual deviations in both variables, *i.e.*, the standard deviations of X and Y. Since both the co-variance and the standard deviations are affected by the size of the variates in absolute terms, by dividing one into the other, this distorting factor is eliminated. If the student reader examines the formula for the coefficient in the light of the foregoing, he may not need to rely on his memory so much.

Calculating r using assumed means

In the second example given below the data are derived from a social survey in London. It is proposed to calculate the coefficient of correlation between poverty and overcrowding. For the purposes of the survey, poverty was defined as living below a prescribed minimum standard, and overcrowding was defined

as living more than two persons per room. These definitions are unimportant for purposes of the calculation. They are, of course, essential for purposes of interpreting the data. The figures for poverty and over-crowding, *i.e.* the variables X and Y, are in the nature of percentages, although they are actually expressed per 200 households. For example, in borough A 17 out of every 200 households were living in poverty, and 36 were overcrowded.

Calculation of co-efficient of correlation between Poverty (defined as living below given minimum standard) and Overcrowding (defined as more than two persons per room) in 12 London boroughs.

Borough	No. per 200 households in Poverty(x)	Over-crowded(y)	(x-x)	(y-y)	(x-x)^2	(y-y)^2	(x-x)(y-y)
A	17	36	+ 7	+14	49	196	+ 98
B	13	46	+ 3	+24	9	576	+ 72
C	15	35	+ 5	+ 13	25	169	+ 65
D	16	24	+ 6	+ 2	36	4	+ 12
E	6	12	— 4	—10	16	100	+ 40
F	11	18	+ 1	— 4	1	16	— 4
G	14	27	+ 4	+ 5	16	25	+ 20
H	9	22	— 1	0	1	0	0
I	7	2	— 3	—20	9	400	+ 60
J	2	8	— 8	—14	64	196	+112
K	10	17	0	— 5	0	25	0
L	5	10	— 5	—12	25	144	+ 60
12	125	257	+ 5	— 7	251	1851	+535

Actual means are 10·42 and 21·42 but to avoid calculations with such awkward figures, we work from assumed means. We select 10 as the mean of x and 22 as the mean of y representing these assumed means in the column headings and in the formula below, as x and y. The formula in this case is as follows:

$$r = \frac{\dfrac{(x—\mathrm{x})\,(y—\mathrm{y})}{N} — \dfrac{\sum(x—\mathrm{x})}{N} \times \dfrac{(y—\mathrm{y})}{N}}{\sigma_{\bar{x}}\,\sigma_{\bar{x}}}$$

$$\sigma_{\bar{x}} = \sqrt{\frac{251}{12} - \left(\frac{5}{12}\right)^2} = 4.5 \quad \sigma_{\bar{y}} = \sqrt{\frac{1851}{12} - \left(\frac{-7}{12}\right)^2} = 12{\cdot}4$$

$$\frac{\dfrac{535}{12} - \left(\dfrac{5}{12}\right)\left(\dfrac{-7}{12}\right)}{4{\cdot}5 \times 12{\cdot}4}$$

$$= \frac{44{\cdot}83}{55{\cdot}8} = +\,0{\cdot}80$$

The data are set out in the same way as in the earlier example, but the calculation is complicated by the fact that the standard deviations and the cross-products are worked not from the true means of the two distributions but from assumed means. In other words, a correction must be introduced. The student will remember the correction required for purposes of calculating the standard deviation. It is shown in the example above. The correction for the sum of the cross-products measured from their true means is quite simple. The cross-products actually given in the final column of this table are based on the deviations measured from assumed means in each distribution. The difference between the true mean and assumed mean for both distributions is given in the working, columns 4 and 5. If the sum of the cross-products in the last column is divided by N, *i.e.* the number of pairs, we obtain the average product. From this we deduct the product of the two errors in the two averages or means divided by N^2. The result is then equal to the cross-products of the deviations as if they had been measured from their true means, and the subsequent arithmetic is as in the earlier example.

The student-reader will not have failed to note the considerable amount of arithmetic required in the first example on page 207. Quite apart from the fact that it contained twenty pairs of observed values, there was disproportionately more calculation than in the second example. As an exercise the student can apply the principles of the second illustration to the first set of data. Note, for example, that all the values of x are rounded to the nearest £10, similarly the y values are rounded to the nearest £5. Instead of taking the true means of x and y, use a multiple of 10 and 5 for x and y respectively. For example, take £350 and £75 as the assumed means turnover and profit. All the deviations of x and y can then be given in multiples of 10 and 5. If, as in the illustration of calculating the standard deviation from an assumed mean in terms of class intervals given on pages 126, you use deviations in terms of the class intervals, the calculations will be much easier. Note that the cross-products will be in units of the product of the two class intervals, *i.e.*, £10 $\times$ 5. If the actual value of the standard deviation for either variable were required, it would be necessary to convert the figure obtained, by multiplying it by the appropriate class interval.

This correction is not required for calculating the coefficient of correlation since the sum of the cross-products making up the numerator in the formula, and the product of the two standard deviations, are both expressed in class-intervals. The answer derived from this calculation will be the same as for the more detailed method. By performing it, the student reader will ensure that he understands the basis of the calculations.

Calculating *r* from grouped data

In the third example, illustrating the calculation of the correlation coefficient the data are set out in what is known as a bivariate table. It will be noted that the pattern of the figures over the grids is somewhat similar in appearance to the scatter diagram discussed earlier. The data in this example relate to the weekly expenditure on accommodation and food of 33 individuals. All figures are given in shillings. It will be noted that instead of the discrete distributions employed in the two earlier examples, this illustration comprises two grouped frequency distributions, one of which is read horizontally, *i.e.* expenditure on food, and the other vertically, *i.e.* expenditure on accommodation. The layout of the calculation should be studied with especial care. The four vertical columns to the side of the table and the four horizontal columns below it are the same except that those to the right show the calculations for the y values and those below for x. The column headed f is nothing more than the frequencies derived by cross-adding the frequencies within each cell. Thus reading from the right-hand columns, there are 9 cases where the expenditure on accommodation is 50s. per week. The second column in each case relates to the deviation from the assumed mean which is 45s. for x and 55s. for y. As before, these assumed means are represented by the symbols x and y, and the third and fourth columns in each case are the sum of the frequencies and deviations and the deviations squared required to compute the difference between the assumed and true mean and the standard deviation. Note at the head of each column the reminder that the deviations are in class-interval units, hence we write $(fy—\mathrm{y}) \div c.i.$

The calculation of the cross-products is more complex, however. As in the two earlier examples, the related pairs of

Expenditure on Accommodation y (shillings)	Expenditure on Food (x) (shillings) 10—	20—	30—	40—	50—	60—	f	y-y ÷ c.i.	$f(y$-y) ÷ c.i.	$f(y$-y)^2 ÷ c.i.
20—					1_{-3}		1	—3	—3	9
30—			1_2				1	—2	—2	4
40—		1_2		2_0	4_{-1}		7	—1	—7	7
50—		5_0		4_0			9	0	0	0
60—			1_{-1}	3_0	2_1	1_2	7	1	7	7
70—				2_0			2	2	4	8
80—								3		
90—	4_{-12}	2_{-8}					6	4	24	96
	4	8	2	11	7	1	33		23	131
x—x ÷ c.i.	—3	—2	—1		1	2				
$f(x$—x) ÷ c.i.	—12	—16	—2		7	2	—21			
$f(x$—x)^2 ÷ c.i.	36	32	2		7	4	81			

$$r = \frac{\dfrac{\Sigma f(x\text{-x})(y\text{-y})}{N} - \dfrac{\Sigma f(x\text{-x}) \times \Sigma f(y\text{-y})}{N}}{\sigma x\ \sigma y}$$

Sum of the cross-products, *i.e.* $\Sigma f(x\text{-x})(y\text{-y})$: $-3 + 2 + 2 - 4 - 1 + 2 + 2 - 48 - 16 = -64$.

$$\frac{\Sigma f(y\text{-y})}{N} = \frac{23}{33} = 0{\cdot}697 \qquad \frac{\Sigma f(x\text{-x})}{N} = \frac{-21}{33} = -0{\cdot}636$$

$$\sigma x = \sqrt{\frac{131}{33} - \left(\frac{23}{33}\right)^2} \qquad \sigma y = \sqrt{\frac{81}{33} - \left(\frac{-21}{33}\right)^2}$$

$$= \sqrt{3{\cdot}969 - 0{\cdot}485} \qquad = \sqrt{2{\cdot}454 - 0{\cdot}405}$$

$$= 1{\cdot}87 \qquad = 1{\cdot}43$$

by substitution we get $\frac{-64}{33} - (0{\cdot}697)(-0{\cdot}636) = -1{\cdot}94 + 0{\cdot}44 = -1\ 5.$

$$r = \frac{-1{\cdot}5}{1{\cdot}43 \times 1{\cdot}87} = \frac{-1{\cdot}5}{2{\cdot}67} = -0{\cdot}56$$

deviations must be multiplied together; but whereas in the earlier examples there was only one of each pair, in the present example there are more than one in several cases. For example, in the bottom left-hand corner of the grid it will be seen that there are four cases in which the expenditure on accommodation is 90s and over, while the expenditure on food is between 10s and 19s. The deviations corresponding to this item are + 4 and — 3. These will be found in the second columns of the calculations beside and below the grid respectively. The product of — 3 and 4, *i.e.* the deviations equal to — 12, is inserted in the corner of the cell containing the four cases. In the adjacent cell, which shows that there are two cases where the expenditure on accommodation ranges from 90s. while expenditure on food is 20s. and over, the appropriate deviations are — 2 and + 4, so that — 8, *i.e.* the product of these deviations, is inserted in the corner of this cell. This is done for each cell which contains a frequency. It will be noted, however, that all the cells opposite the classes containing the assumed means, 40 — in the case of food and 50 — in the case of accommodation, have for obvious reasons a product equal to zero.

The next stage is then to multiply the cross-products of the deviations inserted in these cells by the frequency within that cell, due regard being paid to signs. This is done at the foot of the table, each product being set out individually, yielding in this case a net sum of — 64. Very often the products of the cell frequency and the deviation product are inserted in the cell in the opposite corner to that containing the product of the deviations. This practice can be confusing for the student and entails a double lot of writing since the products have to be summed separately. The student should work these stages through by himself with the example, checking that he fully understands what has been done. The remainder of the calculation is then similar to that in our second example. A correction is required for the fact that the cross-products and standard deviations have both been measured from assumed means. As in the second example, the sum of the cross-products, *i.e.* — 64, is averaged over the 33 pairs, and the product of the differences between the assumed and true means of x and y respectively deducted from the average cross-product. That value is then divided by the products of the two standard deviations.

The Significance of *r*

Correlation analysis has been applied to data from most scientific fields. It has been used to determine the relationship between crop yields and variations in the application of fertilisers; the level of fatstock prices and its relation to the cost of feeding-stuffs. Engineers and chemists employ correlation to determine the extent to which properties of their products are affected by variations in the production processes. Its use has been extended to psychological tests designed to measure aptitude for particular types of work; *e.g.*, accident proneness and temperament. It is largely for this reason, namely the widespread and frequent references to results derived from correlation analysis, that the topic is discussed in this elementary text. In practice, the analysis is complicated by other than the purely statistical problems of technique.

In common with most statistical techniques, correlation analysis is usually employed on samples. Thus r, like other statistics derived from samples, must be examined to see how far the results may be generalised for the population from which the sample was drawn. Significance tests have also been evolved for the correlation coefficient. These lie beyond the scope of an elementary text, not because they are difficult to compute, but simply because, like correlation analysis itself, the technique has to be employed with great care and the interpretation of the data, as well as the results, demands a skill and knowledge of the field of enquiry only possessed by the expert. The value of correlation analysis is underlined by the variety of fields in which it finds application, but at all times it is essential to consider the data and ask 'what is the nature of the relationship measured by the coefficient?'

In the illustrative examples given above, the samples were extremely small, although in our final example, which contained 33 paired observations, it may be conceded in theory at least that this is a large sample. The most difficult problem is to interpret the value of the correlation coefficient. Thus the fact that in our final example the correlation coefficient was $-0{\cdot}5$ might, since r falls considerably short of unity, lead the reader to assume that the correlation between the two types of outlay is negligible. Unfortunately it is not possible to interpret the correlation in this way. It is not possible, for example, to say that a

value of r equal to 0·9 is very high and more significant than one equal to 0·8 since it will depend upon the size of the sample used. Nor should too much be read into the coefficient. In the first example, which showed the relationship between the turnover and profit margin, it is apparent and reasonable to assume that these two variables are interrelated, and that the margin is presumably dependent upon turnover, *i.e.* the relationship is causal. The coefficient of correlation tells us nothing about the nature of the relationship; it merely indicates its existence. It is for the statistician to interpret it and deduce its nature and significance. It is in this respect that regression is so useful, since it defines in exact terms the relationship between the two variables. In the second example, the correlation coefficient of + 0·8 suggests a significant relationship between poverty and overcrowding which is probably causal too, *i.e.* people are overcrowded because they are poor. At all times, however, one must beware of drawing dogmatic conclusions from limited data.

One final use of correlation analysis may be mentioned. Generally speaking, the square of the value of r may be regarded as the percentage of the variation in Y directly attributable to changes in X. Thus, as far as the first illustration is concerned, approximately 81 per cent of the variation in Y is explained by variations in X. This figure is known as the 'explained variance', while the balance of 19 per cent is termed the 'unexplained variance'. This means that as far as the available data are concerned, no precise explanation of the cause of 19 per cent of the variation is given. It may be attributable to any or many of several causes.[1]

Rank Correlation

The methods of correlation that have so far been demonstrated have all been concerned with the measurement of the relationship between series of *numerical values*. It is possible, however, to measure the degree of correlation between two sets of observations or between paired values when only the *relative* order of magnitude is available for each series. For example, suppose a group of students sat two papers in an examination and instead of the actual marks awarded on each paper they

[1] For the interested reader with some knowledge of algebra, the H.M.S.O. publication 'Industrial Experimentation', 4th Edition by K. A. Brownlee, provides a full account of correlation analysis and variance analysis in industrial research.

were told only their ranking in order of merit. If it was desired to establish whether the performances on the two papers were correlated or not the method of *rank* correlation could be used.

The coefficient of rank correlation is given by Spearman's formula:

$$r_r = 1 - \left(\frac{6\Sigma d^2}{n(n^2-1)}\right)$$

where d is the numerical difference between corresponding pairs of ranks and n the number of pairs. In the following example, ten students are ranked in order of merit on two examination papers in French and Latin.

Student	Rank in French	Rank in Latin	d	d^2
A	1	3	2	4
B	2	2	0	0
C	3	1	2	4
D	4	6	2	4
E	5	5	0	0
F	6	8	2	4
G	7	4	3	9
H	8	10	2	4
I	9	7	2	4
J	10	9	1	1
				34

Substituting the values derived from the above table in Spearman's formula

$$r_r = 1 - \left(\frac{6 \times 34}{10(10^2-1)}\right) = 1 - \left(\frac{204}{990}\right) = +\ 0{\cdot}79$$

which suggests quite a strong relationship between performance in the two papers.

As well as in the type of problem just illustrated the coefficient of rank correlation may be calculated for series of *qualitative* instead of *quantitative* data, *e.g.*, colour of hair and intensity of emotion measured on a non-numerical scale, or any attribute which cannot be measured numerically. Similarly we could calculate the coefficient of rank correlation for any group of paired observations even if they were numerical values, *e.g.*, marks instead of placings in an examination, and the normal coefficient of correlation could be calculated by the product

moment formula shown earlier. Significance tests for rank correlation do exist, but they are not relevant in this elementary text.

As with all the techniques described so far, correlation analysis has no value for its own sake. It is useful solely because, if properly used, it permits theories and hypotheses to be verified or rejected on the basis of empirical evidence. At all times it must be remembered that such specialised tools may easily be misapplied and give misleading results.

NOTE TO CHAPTER XII

Calculation of Regression Lines

The line of regression of y upon x is given by the equation $y = a + bx$ if the relationship between y and x is linear. We need to determine the value of the two constants a and b. It is known that when x has its average value, the best estimate of the corresponding value of y is its own mean. Thus, we have two values for the above equation, $\bar{x} = 363$ and $\bar{y} = 72$.

The value for b, which gives the slope of the regression line, is derived from the equation $b = \frac{\Sigma xy}{\Sigma x^2}$ where x and y are the deviations of individual values of x and y from their respective means.[1] The numerator requires the deviations for each pair of x and y values to be multiplied together and their products added. This is done in col. 6 of the table below. The denominator in the above equation is the sum of all the squared deviations of x values from their mean – as given in column 5. This is done in the same way as if we were proposing to calculate the standard deviation of the series of x values.

Thus if $b = \frac{\Sigma xy}{\Sigma x^2}$ then $b = \frac{53,780}{254,220} = 0{\cdot}2115.$

We now have three of the four values in the equation $y = a + bx$, *i.e.*, $x = 363$, $y = 72$ and $b = 0{\cdot}2115$. The fourth value a is easily derived:

$72 = a + 0{\cdot}2115(363)$

$72 = a + 76{\cdot}8 \therefore a = -4{\cdot}8$

Therefore, the complete regression equation of y on x as drawn in Figure 19B reads: $y' = -4{\cdot}8 + 0{\cdot}2115\ (x)$. The y'

[1] Or in the alternative notation at the top of the columns on page 207, $\frac{\Sigma(x - \bar{x})(y - \bar{y})}{\Sigma(x - \bar{x})^2}$

symbol indicates that this value of y is a calculated or predicted value, using this equation.

We can test the equation by substituting some of the known values of x, *e.g.* 200, and compare the predicted value of y, which is 37·5 when $x = 200$, compared with an observed value of 35. Thus the correspondence is quite close, as we would expect when the value of r is so high. It is possible to calculate the standard deviation of the differences between the observed and predicted values of y for given values of x, and thereby estimate the standard error, *i.e.*, the extent to which such differences arise from random sampling errors, but this lies beyond the scope of our present study.

Just as we have calculated the regression line of y upon x, we can calculate the line of x upon y. In this case, the value of b is given by the equation $b = \frac{\Sigma\, xy}{\Sigma\, y^2}$. The calculation of Σy^2 is given in column 7 below and the student can repeat the calculations given above to obtain the regression equation of x upon y. The answer is $x = 36{\cdot}1 + 4{\cdot}53y$.

x (1)	y (2)	$(x - \bar{x})$ (3)	$(y - \bar{y})$ (4)	$(x - \bar{x})^2$ (5)	$(x - \bar{x}) \times (y - \bar{y})$ (6)	$(y - \bar{y})^2$ (7)
150	30	—213	— 42	45,369	8,946	1,764
200	35	—163	— 37	26,569	6,031	1,369
210	40	—153	— 32	23,409	4,896	1,024
230	45	—133	— 27	17,689	3,591	729
260	50	—103	— 22	10,609	2,266	484
280	50	— 83	— 22	6,889	1,826	484
300	60	— 63	— 12	3,969	756	144
320	65	— 43	— 7	1,849	301	49
350	70	— 13	— 2	169	26	4
370	70	+ 7	— 2	49	— 14	4
380	80	+ 17	+ 8	289	136	64
400	75	+ 37	+ 3	1,369	111	9
410	85	+ 47	+ 13	2,209	611	169
430	90	+ 67	+ 18	4,489	1,206	324
460	100	+ 97	+ 28	9,409	2,716	784
470	80	+107	+ 8	11,449	856	64
480	90	+117	+ 18	13,689	2,106	324
500	100	+137	+ 28	18,769	3,836	784
520	110	+157	+ 38	24,649	5,966	1,444
540	115	+177	+ 43	31,329	7,611	1,849
$\bar{x} = 363$	$\bar{y} = 72$	0	0	254,220	53,780	11,870

In the foregoing example it was possible to work out the value of b from the true means of x and y. This is not always the case

and just as when calculating the mean or the standard deviation we may have to work from assumed means, so it is possible to derive the regression equations by working from assumed means. In such a case, the formula for b for the line of y on x reads

$$\frac{\Sigma\, xy - \frac{\Sigma\,(x)\,(y)}{N}}{\Sigma\, x^2 - \frac{(\Sigma\, x^2)}{N}}$$

The corrections $\frac{\Sigma\,(x)\,(y)}{N}$ and $\frac{(\Sigma\, x^2)}{N}$ are really the same as we should use when calculating the standard deviations of x and y. Space does not allow us to show all the details, but if the data given in the table were re-calculated taking as the assumed mean of x 350 and of y 75, then the total line at the bottom of the table would read as follows:

$\Sigma(x-\mathrm{x})$	$\Sigma(y-\mathrm{y})$	$\Sigma(x-\mathrm{x})^2$	$\Sigma(x-\mathrm{x})\,(y-\mathrm{y})$	$\Sigma(y-\mathrm{y})^2$
+ 260	— 60	257,600	53,000	12,050

Substituting in the above equation for b, we get:

$$b = \frac{53{,}000 - \frac{(260)\,(-60)}{20}}{257{,}600 - \frac{(260)^2}{20}} = 0{\cdot}2115$$

Since we are working from assumed means a further correction is required to derive the regression equation:

$$\begin{aligned} y - \bar{y} &= b(x - \bar{x}) \\ y - 72 &= 0{\cdot}2115(x - 363) \\ y &= 72 + 0{\cdot}2115x - 76{\cdot}8 \\ y &= -4{\cdot}8 + 0{\cdot}2115x \text{ (as on p. 219.)} \end{aligned}$$

[1] As before we are using x and y to represent individual values of each variable, with x and y representing the assumed means of the two variables.

CHAPTER XIII

SAMPLE SURVEYS

Definition

To the layman a sample survey may appear to be an inferior substitute for a census. Reasons for not taking a census may be that the population in question is too large and the census would take too long and would be too expensive. In some cases, however, the sample survey may be preferable to the census, not merely on account of the lower cost and greater speed with which results are made available, but because it is superior to the census for the particular enquiry. The Director of the Social Survey has defined a sample survey as a *method of collecting detailed information relating to representative groups under controlled conditions.*[1] This particular definition brings out all its main features. The outstanding advantage of such a survey over the census lies in the fact that it is practicable to collect much more detailed information from a relatively small number of people than from a large number. The former method permits the use of trained interviewers who can elicit a great deal of detailed information not merely of a factual nature, but also opinions. The census is only satisfactory for collecting factual data and even in that case some of the information received must be regarded as distinctly doubtful in terms of its accuracy.

The definition emphasises that the sample is representative, a fact which is all too often taken for granted. In theory, however representative it may be, it is still inferior to the 100 per cent enumeration of all the population units which is implied in the term 'census'. In practice, however, no census covering a population of any size at all, is 100 per cent complete. For example, as was cited earlier, the Register of Electors in this country appears to have a deficiency of 4 per cent due to non-registration. The U.S. Bureau of the Census has carried out a series of investigaitons to test the reliability of data assembled from its 1950 population census which reveal quite substantial 'under-counts'

[1] The Scope of Sample Surveys, L. Moss. Read before the Conference on Modern Sample Survey Methods, organised by the Association of Incorporated Statisticians, December 1953. This, together with the other papers read before the conference, has been published by the Association in booklet form.

and deficiencies for certain groups within the population. Clearly this type of problem is much more serious in a country so great in area and so diversified in race, education and language as the United States. But to a lesser degree the problem is present in any census. Lastly, the survey has the merit that the information is collected under what Mr Moss has called 'controlled' conditions. It is difficult, he states, 'to exercise close control at the critical point, the point when the informant's information is put down on the questionnaire or schedule.' Even when the census authorities use enumerators (few of whom can be really well trained in the work) to collect the forms and where necessary help the respondent to complete them, this weakness is very serious in any large census. With a sample survey covering at most a few thousand informants, it becomes possible to employ skilled investigators. They can explain the purpose of the survey to the respondent and ensure that each question is correctly put and understood.

It is for such reasons that the sample survey has come into prominence for its own merits rather than as an inferior substitute for the infrequent and cumbersome census. As Mr Moss has pointed out, 'experience seems to show that it is wrong to assume that a census must automatically be more correct than a sample'. Indeed, American experience suggests that the only method for testing the reliability of census data is a properly designed sample survey!

Development of Surveys

The enumeration of populations by means of a census is centuries old; the Egyptians and Romans both carried out censuses for fiscal and military purposes. The sample survey, however, is of quite recent origin. It is, nevertheless, customary to start any history thereof with the great social enquiry of Charles Booth entitled 'Labour and Life of the People of London' which filled 17 volumes and took more than a decade to complete. A few years after Booth had published his main findings, Seebohm Rowntree carried out his enquiry into poverty among the working classes of York. This was published in 1901 under the title 'Poverty: A Study in Town Life'. Neither of these pioneer enquiries could be described as sample surveys. They were virtual censuses of the relevant populations. In

Booth's case the population was the working class family with children of school age. For Rowntree it was all working class families, the latter being defined as households where no domestic help was kept!

The first survey based upon a random sample of the population was carried out in 1912 by the late Professor Bowley in Reading. Like his great predecessors, he was concerned to measure the incidence of poverty among the working classes. He incorporated Rowntree's device of measuring the incidence of poverty by reference to a minimum living standard, but in place of a census used a one in twenty sample of addresses taken systematically from a street directory. With the growth of unemployment after the war, surveys into poverty were undertaken in many cities. The best known are those in London by Professor Bowley and his assistants from the London School of Economics and that on Merseyside prepared by Caradog Jones of the University of Liverpool. It was not until the later 1930's that the public became at all aware of the growing use of sample surveys. Their attention to this subject was attracted by the well-publicised public opinion polls which had established a considerable reputation in the United States. One or two commercial agencies were also beginning to adapt the technique for market and consumer research purposes. On the outbreak of the war, the Central Office of Information created the Wartime Social Survey.

The Social Survey, as it is now called, is still in being and is contributing much to the development of the techniques used in sample survey work. Originally it was employed by various government departments on *ad hoc* surveys to learn what the public felt about certain issues, *e.g.*, clothes and fuel rationing. After the war, it carried out surveys into labour problems, *e.g.* why recruitment to the nursing profession was so poor and why miners were leaving the mines. More recently it assisted with the survey of savings carried out by the Oxford Institute of Statistics and the large-scale Ministry of Labour Enquiry into Household Expenditure which provided the basis for the new cost of living index. Many of the reports prepared by the Social Survey, particularly those of recent years, are confidential and have not been published. Those that are available in libraries, however, are models of their kind and the student cannot do better than

study them if he wishes to understand survey techniques. Of particular value are the periodic articles and papers from members of the Social Survey's staff which have appeared in the learned journals.[1]

Among other well-known survey bodies are the B.B.C. Audience Research Department and the British Institute of Public Opinion, better known to newspaper readers as the 'Gall-up Poll'. Both these bodies use quota sampling in their surveys. The B.B.C. Audience Research uses part-time interviewers for ascertaining public reaction to both the T.V. and sound broadcasts. Approximately 3,000 people in the seven B.B.C. regions are interviewed each day and asked what programmes they heard or viewed the previous day. The results are not made public but are circulated within the Corporation in the 'Audience Barometer' for the benefit of programme advisers and others. In addition to the field work by interviewers, who only work for about three weeks at any one time since the work tends to become both tiring and boring after a while, the B.B.C. maintains separate panels for sound and T.V. comprising nearly 5,000 households and individuals. The composition of the panels reflects the socio-economic structure of the population. There is no difficulty in obtaining new recruits for these panels, either sound or T.V. Neither panel, however, can be regarded as fully representative of the listening or viewing public, since only those really interested in viewing or listening will take the trouble to volunteer and provide the information required. There is also the problem common to any panel, that its members become 'conditioned' and cease to be representative of the general population.

Stages in the Survey

Most surveys are designed to assess individual views on current political economic and social issues; they have in common a carefully designed schedule of questions and intensive interviewing. Different types of schedule ~~will be~~ employed for various types of enquiry, just as the sample may differ. The basic pattern, however, of all surveys may be described quite briefly. Without some understanding of the principles and problems of survey work an adequate assessment of survey data and results is hardly

[1] Some references are given in this chapter.

possible. As a start it is sufficient merely to detail the successive stages in such a survey.

1. The first stage may be described as providing an answer to the question 'what is the problem under review and how may the survey help?' This cannot be answered without a detailed study of all the facts. The organiser may have to spend a long period immersing himself in the subject and learning just what his client's needs really are. The maximum information must be derived for a given cost. Not merely must the information be relevant to the enquiry but it is important to ~~avoid~~ the situation which can so easily ~~arise at~~ the end of a survey, when it becomes clear that it would have been helpful if only some additional data had been collected on a particular point. All this can only be achieved by the most careful consideration of all the facts; too much time cannot be spent on these initial stages if time and money are to be spent to the best advantage. Such observations may appear somewhat trite and obvious; it is nevertheless surprising how often it is the obvious which is overlooked!
2. How is the information to be obtained? The two main methods are the postal enquiry and the survey employing interviewers. The merits of each are discussed below.
3. The preparation of the schedule of questions and instructions for their completion. A badly designed questionnaire may ruin an otherwise well conducted survey. If machine tabulation is employed the answers to the individual questions in the schedules will have to be coded. It may seem premature to discuss tabulation at this stage, but the main tabulations, particularly those which are to bring out the inter-relationships between different characteristics of the sample units, should be carefully prepared at the outset. This ensures that there is no danger of omitting to ask for information which will be needed. Information is sometimes required from a particular section or sections of the sample which may contain relatively few units. With so small a sub-sample it would be impossible to obtain the necessary degree of precision in the results. The size of the sample will have to be increased or special steps taken to increase the number of units in the particular sub-sample. Such a point could easily be

overlooked in the initial stages of the survey unless all the analyses of the final data are considered in advance.

4. The sample selected must be of such size and composition that it will yield the most reliable results for a given expenditure. If interviewers are to be employed, where the sampling is random or systematic, 'substitutes' may sometimes be drawn in the same way. If quota sampling is to be used, suitable quota sheets and instructions for classifying respondents are necessary.
5. The preparations having proceeded so far it may be advisable to test the schedule of questions by a *pilot* survey, which is a survey in miniature. The number interviewed is unimportant, but the respondents are selected at random. Usually the more experienced interviewers are engaged on this pilot survey, since they are capable of assessing the weaknesses of the approach or any questions on the schedule. The results are not so important in themselves as are the lessons learned. It is not always possible to carry out a pilot survey, although it is desirable. The reasons for its omission may be financial or, more often, simply a question of time. Usually, if the organisation has had a wide experience of similar surveys, *e.g.*, the same survey was carried out, say, eighteen months ago, then clearly the pilot survey may be dispensed with.
6. Before the field work starts a briefing conference for the interviewers is normally held. Any difficulties met and the lessons learned in the pilot survey are examined and a course of action laid down for specified circumstances. The interviewers will have been issued with their instructions and care should be taken to ensure that they fully understand them.
7. Soon after the field work begins, completed schedules will begin to pour into the office. The schedules should be edited for omissions and any obvious mistakes such as inconsistencies in the replies entered. When necessary, the area organisers may check back on the respondent.
8. If the information is transferred to punched cards, the tabulation may be rapidly completed by machine. The questions should have been coded in advance for this purpose. The classification of answers which are not of the simple 'Yes/No' or 'once a week/more often' variety, but may be expressions

of opinion, will require careful consideration and supervision. The data once assembled, the report on the survey may be prepared. Usually several people will discuss the results together to ensure that all aspects of the information are brought out and correctly interpreted.

These, then, are the main stages of any survey. Because each stage has been dealt with separately, it should not be imagined that each is independent of the others. Each survey must be considered in its entirety. At every stage, what has gone before, or what is to be done later, must influence the design of the survey. For example, the type of information sought will largely decide whether the postal or personal enquiry method is used, while the type of respondent will influence the design of the questionnaire.

Problems can arise all along the line, but if a survey is really well planned many difficulties may be anticipated and provision made accordingly. For example, interviewers may be required to classify their respondents by social class. Unless a method of classification suitable for the survey is determined in advance and the interviewers instructed in its application, part of the data may be valueless. Different interviewers will assess individual respondents by different standards, with obvious results. It was precisely this type of problem which has led to the loss of much useful information from the National Farm Survey carried out in 1942. Part of the schedule required an evaluation of the quality of the farm holding by layout, type of farming and condition of buildings. Unfortunately, the interviewers available were inexperienced in survey work. They were normally employed by the County War Agricultural Committees and were usually local men. Consequently, they assessed the holdings for the purposes of these questions in the light of their local knowledge. The result has been that the data are most unreliable for inter-county comparisons, although they are probably satisfactory for providing a local view of farming in the individual counties.[1]

The quality of the data derived from any survey rests largely on the efficiency of the work at three stages. The first is the selection of the sample. If this is unsatisfactory, then clearly no reliable conclusions may be drawn about the population from which the sample was taken. Secondly, the design of the schedule

[1] *See* National Farm Survey of England and Wales, H.M.S.O., 1946, para. 8. A copy of the schedule used in that enquiry is reproduced on p. 30

of questions requires careful thought. Since few respondents appreciate the full implications of any lengthy question, the questions must be such that only one interpretation is possible. Finally, the all-important task of interviewing. As will be seen, more mistakes may creep into the results at this stage than any other. The remainder of this chapter will be devoted to these three basic problems in survey work.

The Sample

Generally speaking the object of a sample survey is to learn something about the population. If the sample is truly random then, as has been pointed out earlier, certain conclusions regarding the population may be inferred from the evidence of the sample statistics. If the sample is, for whatever reason, unrepresentative of the population from which it was drawn, then the sample results cannot be generalised with any confidence to the population. Mr Moss in his paper quoted above, gave two illustrations of the significance of ensuring that the sample is representative. In 1946 a government committee studying the problem of shop closing hours requested a sample survey to assess public opinion on the matter. Evidence already submitted by interested parties had suggested that a certain change would affect only a minority of the public. The survey revealed that this was true, but it also revealed that this minority comprised mainly working housewives who at the time constituted an important part of the labour force. Without the survey this highly important piece of information would never have come to the notice of the committee. In the United States much publicity has been accorded to Dr Kinsey's studies of the sexual behaviour of human beings, in particular the recent report on the human female. As Dr Kinsey himself has emphasised, the sample of informants was for obvious reasons largely self-selected, *i.e.* volunteers only. When the composition of the sample is compared with the entire American female population over 14 years of age, it emerges that Dr Kinsey's sample is seriously overweighted with the younger married woman who has had the benefits of a college education. In other words, the report *may* be a fair summary of the behaviour of this particular group of American women, but no inferences regarding the female population at large can be drawn from it. As is so often the case

when a sample survey attracts publicity, the warnings of the organisers tend to be overlooked.

One of the weaknesses of random sampling is the fact that not every member of the sample can be contacted, or if contacted by the interviewer may not be willing to answer the questions put to him. This question of 'non-response' is met in quota samples by the simple expedient of seeking out additional suitable respondents, *i.e.*, those who correspond to the interviewer's list or quota of respondents, until such time as the interviewer's quota is complete. This policy does not really solve the question of 'non-response'; it merely ensures that the same number of interviews are made as were intended when the sample was designed. The 'non-cooperating' members of the public whom the interviewer sought to interview may, however, form a particular group which will as a result of non-response be under-represented in the sample. If this is the case, then the sample is biased or incomplete and the results of the survey to that extent cannot reflect the true position. The problem of non-response with random samples in which the interviewer is given a list of names and addresses is at all times serious. The experience of the Social Survey is that in a survey of adults the chances that an interviewer will find the respondent at home on her first call are about one in three. If the respondent is not available, then a further visit is necessary. Experience shows that a maximum of three calls is the economic limit. Admittedly, continuous attempts to establish contact will produce a larger proportion of effective interviews out of the sample, but the improved results are of disproportionate value to the efforts involved. Hence, the Social Survey instruct their interviewers to make a maximum of three calls at any one address to find the prospective respondent. Until a few years ago the Social Survey interviewers were given lists of 'random' substitutes in the event that contact with any address in their first list of interviews was impossible; for example, because the person had either died or moved out of the district. This practice of substitution has long been abandoned, for as was pointed out above with quota interviewing, its only merit is to ensure that the interviewer carries out the required *number* of interviews. Nowadays the emphasis is placed on the need to obtain a satisfactory interview with the selected individual. No substitutes are provided, but the interviewer is

required to make a note of all failures and unsatisfactory interviews so that these 'non-respondents' can to some extent be classified. The object of such a policy is to endeavour to judge whether the non-respondents are merely a sub-sample of the main sample, in which case the only drawback is that the statistician will have a smaller sample than he had hoped for, and his results will be to that extent less precise.

The real danger is that the non-respondents will form a particular group which in consequence of the non-response will be under-represented in the sample. If some means of classifying these non-respondents can be found, then this risk can be reduced. For example, suppose a sample of households has been interviewed and it appears that 40 per cent of the households have no children in them. It is known from the census data that 57 per cent of all households are childless and it is clear that this sample is deficient in such households. Therefore the views of such households will not be given their due weight in the final analysis. The probable explanation of this deficiency is that in childless households all the adults go out to work and when the interviewers called they received no reply. They may have been slack about call-backs and in consequence an insufficient number of such households have been interviewed. It is for such reasons that classificatory questions are introduced into the schedule; for example, the number of children, size of income, occupation, daily newspaper read, among others.[1] The distribution of the population in respect of certain such characteristics is known and the sample should correspond with the population.

The value of such classificatory questions is largely attributable to the wealth of information collected in the decennial census of population. These data can be supplemented and in some cases kept up to date by reference to the annual reviews of the Registrar General.[2] Thus the conventional classificatory questions relate to age, sex and region, the replies to which can then be compared with the official data for the population as a whole. In the I.P.A. Readership Survey the replies to a number of other questions in the schedule, which were of interest in themselves, could also be used to test whether or not the sample was representative. Thus, replies to a question asking at what age the

[1] See the classification schedule abbreviated on page 245 and the opening section of the schedule on p. 238.

[2] See Chapter XV.

respondent's full-time education ceased were compared with the information derived from a similar question in the 1951 census of population. Additional questions relating to the ownership of a car, a T.V. set, telephone and refrigerator also provided further checks upon the sample. Car ownership admitted or claimed could be verified against the national figures prepared by the Ministry of Transport; while cinema attendances are compared with the data published each quarter by the Board of Trade. The check data in respect of T.V. ownership were provided by the results of an earlier enquiry carried out by the B.B.C. Audience Research department, and that relating to the ownership of refrigerator by a survey conducted by the Odhams Press research unit. Obviously there is a limit to the number of classificatory questions that can be inserted in a schedule but the more 'control' questions that there are, the better. American research has shown that apparently satisfactory results sometimes emerge when a sample is compared with one type of control, but when other control data are used, the sample is deficient. In other words, the more cross-checks on the sample composition the better, and it is desirable that the checks themselves should be independent of one another.

An interesting illustration of measuring the extent of non-response and making allowance for it is provided by the 1946 Family Census. This enquiry covered a sample of over 1·7 million married women. When the returns were checked there was a deficiency of some 17 per cent and it was suspected that among this group, childless women were in the majority. This was an especially important group in what was really a study of human fertility. Follow-up letters asking the 'non-respondents' to co-operate produced a proportion of replies and from these it was clear that the suspicion was fully justified. In the results, the figures for childless married women were adjusted in the light of this knowledge.

The control and measurement of non-response remains among the more intractable problems of survey organisers. The solution is to be found partly in first-class interviewing with well-designed schedules of reasonable length and partly in checking, as above, on the non-respondents so that allowance for their omission from the sample can be made. This should not be read as implying that the organisers guess the facts about

them. It means that the answers given by respondents who appear to be similar – as far as they can be compared by reference to certain classificatory data – can be proportionately weighted in the final analysis. Because of the danger that non-response may introduce bias into the sample, it is far better to use a smaller sample in which interviewers are expert and can ensure accurate replies as well as a very high response rate, rather than a much larger sample with poorer interviewing and a lower response rate. Even if the latter sample yields a larger number of interviews, the bias may lead to erroneous conclusions.

The results of two important surveys in respect of the proportion of the sample successfully interviewed will illustrate the type of problem that the organisers have to deal with. In the 1957 I.P.A. national readership survey which was based upon a random sample of individuals whose names and addresses were taken from the Electoral Register, the proportion of successful interviews was 78 per cent of the original sample drawn. This consisted of 12,160 names but of this number 1,332 had either died, moved away, or the premises at the address were empty or demolished. Effectively the interviewers had to contend with 10,828 available respondents and of this number another 1,339 proved to be failures. Almost half of them refused to be interviewed, one-third of them were out on each of three or more calls and about one-sixth of their number were either sick, senile or otherwise un-interviewable. The effective sample of 78 per cent was analysed by regions and by age, the results being weighted to adjust for any under-representation.[1] The other survey is of interest because it followed what might be an extremely unwise procedure. In the Household Expenditure Enquiry of 1953/4 the information required of all members of households was a detailed analysis of their expenditure over a three-week period. From experience with the recent National Food Surveys it was anticipated that the refusal rate in the sample would be high; a rate of 60 per cent was considered likely. For purposes of the survey it was considered that a sample of 8,000 effective interviews and completed budgets would suffice, and on the basis of an expected 40 per cent response some 20,000 households were selected as the sample. In the event, about 65 per cent, *i.e.*, some

[1] Both the 1957 and 1958 editions of the I.P.A. National Readership Surveys are described in extremely clear and detailed explanatory memoranda which set out the sample, schedule of questions and interviewer instructions.

13,000 households co-operated. The important point to note is that the organisers of the survey would have to check the composition of their effective sample of replies very carefully against the known make-up of the population to ensure that it was fully representative. Furthermore, those households which co-operated might have rather different expenditure and consumption habits from those who refused to cooperate. In the event, the Cost of Living Advisory Committee which was responsible for the survey declared that the sample of some 12,900 returns, which they had used for constructing the new Index of Retail Prices, could be regarded as fully representative of households in this country.[1]

Designing the Questionnaire

The questionnaire or schedule of questions may be described as the keystone of any survey. The basic problem is not so much what questions to ask, but what is the best way to ask them. According to two members of the Social Survey, 'the problem is . . . to design questions that mean the same thing, a single thing, a defined thing and the intended thing, to everyone'.[2] As a general rule, questions should be short; lengthy questions tend merely to confuse the respondent. The interviewers should be instructed as to whether they may depart from the form of the question as written on the schedule. Usually, the question has been carefully considered and the final form is probably the best possible whereas the interviewer's alternative may lead to erroneous interpretation of its meaning. An obvious case is to avoid mentioning proprietary brands when engaged on a market survey. The replies to the question 'What do you consider the best wireless set costing less than £20?' will be very different in the aggregate from those received if the interviewer had asked, 'Do you consider the —— radio the best set costing less than £20?' The mere mention of a name, or the hint of the exact purpose of the enquiry, is generally sufficient to influence the respondent. Even the interviewers should not know which organisation is paying for a survey to determine consumer preferences among such products as detergents, newspapers or soft drinks.

[1] For an interesting account of this survey see 'The Report of an Enquiry into Household Expenditure in 1953/4, prepared by the Ministry of Labour and published by H.M. Stationery Office, 1957

[2] Fothergill and Willcock, 'Interviewers and Interviewing' in Modern Sample Survey Methods.

Only questions which the respondent can answer from knowledge or experience should be asked. To ask a rural housewife who has always used either open fires or paraffin stoves if she prefers to cook by gas or electricity is pointless. Yet in one survey housewives were asked which form of heating they preferred, coal fires, gas, electricity, or central heating. Not many housewives in this country have enough experience of the last-mentioned form of heating to be able to answer the question rationally.

A particular problem in interviewing is the respondent's memory. Too much reliance should not be placed on it! In consumer surveys the interviewer, instead of enquiring about the consumption of a product over a period, *e.g.*, a month, usually asks whether the housewife has a particular commodity in the house, *e.g.*, a soap powder, and if so, when she bought it. Alternatively, they may be asked how much of a foodstuff, *e.g.*, biscuits, they bought in the current week. In the Hulton Readership Survey, informants were asked which newspapers and periodicals they read. To avoid the risk of any paper being overlooked, each informant was shown a short list and asked to indicate those that he had seen. The same technique is used by the interviewers of the B.B.C. Audience Research. In this case the previous day's broadcasts are listed. Even this device is not perfect; experiments have revealed that the position of any item is important, those at the top being mentioned more frequently than others. To overcome this difficulty the lists are usually rearranged at short intervals. Nor is the facile assumption that the respondent will answer accurately even the simplest question justified by experience. The number of wives who do not know their husband's incomes is legion, but one consumer survey in a North London suburb revealed that quite a number either had no knowledge of their husband's occupation or described them incorrectly. The same survey revealed a surprising number of housewives who were unaware of the fact that most wireless sets were run off the mains electricity.

All the questions asked must be so phrased that the respondent can answer them intelligently. This implies that the respondent must understand what is being asked. For this reason, in many opinion polls, a factual question or two is usually inserted at the beginning to find out whether or not the respondent

knows anything about the subject upon which he is asked to express an opinion. Every question which relates to the respondent's actions in the past must be most carefully considered, because the average person's memory is so unreliable. Great care should be taken to avoid words with 'emotional' content, *e.g.* 'Socialist' and 'Tory' will probably affect the respondent more than (say) 'Labour' and 'Conservative'. A good illustration of this was provided by a U.S. public opinion poll during World War II. Many more respondents were 'anti-Nazi' than were 'anti-German' when asked their views on the belligerents before December 1941 in two polls, the only difference between which was the use of the word 'Nazis' in place of 'Germans' in the relevant questions.

Apart from basic principles to be observed in the construction of schedules and questionnaires touched upon above, there is also the problem of facilitating the work of the interviewer. A good deal of study has been given to the best lay-out of the form, bearing in mind that it may have to be completed on the doorstep and not on a table. Instructions to the interviewer must be set in bold type, *e.g.*, if respondent answers 'No' to this question omit next section. It is both impracticable and undesirable that the interviewer should have to try and write down the respondent's answers each time. For many questions the answers can be anticipated, *e.g.*, Yes/No/D.K., *i.e.*, respondent doesn't know. Similarly, where the frequency of a particular event is concerned such as the number of weekly visists to the cinema, the answer can be pre-coded, once/twice/more often/ seldom/never. The pre-coding is not always quite so simple. For example, take the question 'Would you say that television has made your home life more interesting and happier, or do you think your family life would be better without it, or does it make no difference'. This question set in a survey on I.T.V. carried the following pre-coded answers: More interesting/better without/ no difference/ don't know. The weakness of this type of pre-classification is that it forces the interviewer to classify the respondent's answer on the spot. Suppose the reply is on something like the following lines 'well, it keeps the kids happy and gives the missus a chance to put her feet up while they are viewing; but sometimes we have trouble in getting the kids to do their homework and I think that the missus is too soft in letting them

stay up at night after their proper bed-time; but it's O.K. in the evenings sitting with the missus after supper and not having to go out . . . etc. etc.' How would the reader classify this reply?[1] Another point to observe when formulating opinion questions is that the respondent should not be given a 'middle course' to follow. For example, 'Do you think that unemployment in this country is likely to rise or fall during the next few months, or do you think that there will be little change'. Since this question involves making an assessment of the future and quite apart from the fact that most respondents will not have the knowledge to form an opinion, it is logical for a large number of the respondents to opt for the middle course. This is particularly the case since it comes at the end of the question and obviously offers them an escape from the difficult choice posed by the earlier part of the question. Such questions which offer the respondent certain alternatives are known as *dichotomous* questions (if there are two alternatives) and multiple choice or *cafeteria* questions where there may be several alternatives. Sometimes an attempt is made to pre-code all the possible answers, the interviewer ticking that one which most closely resembles the respondent's reply. A good example of this is given in Question 11 of the schedule on Diphtheria Immunisation reproduced below. Quite often, in order to create a feeling of confidence in the respondent, the interview will start with what is termed a free-answer question. This is of the variety, 'The government has been asked by a section of the public to re-introduce flogging for crimes of violence. What do you think?' The respondent can then speak quite freely – always assuming he has some views on the subject – and he is then ready for more specific questions. Sometimes these 'free-answer questions' appear later in the schedule, but whenever they appear the interviewer is confronted with the problem of summarising accurately but concisely the gist of the reply. This is not always easy and when the forms go in for tabulation, the classification of such free answers is inevitably arbitrary. While the drafting of a schedule of questions may at first sight appear quite a simple task – after all, what is difficult about asking questions? – in practice it is a highly skilled task. The reader should examine the schedules reproduced in various reports and assess them in the light of the above comments.

[1] This question is taken from the schedule in a report 'Parents, Children and Television' published for the I.T.V. Authority and available from H.M. Stationery Office, 1958.

The length of the interview will vary considerably from one survey to another, but it is a good rule to try and keep it short. The closer the subject of the survey to the respondent's experience and daily life, the longer can be the interview without producing weariness or plain boredom. Nevertheless, the Social Survey appears to have little difficulty in persuading its respondents to withstand interviews of an hour or more in some cases Some schedules contain over 60 questions, but a short questionnaire used in the Diphtheria Inquiry of 1946 is given below. This illustrates the main features discussed so far. All the questions are simple and to the point. The interviewer is helped by the precoding of the probable answers to some questions, *e.g.*, numbers 11 and 12. With this questionnaire is published a detailed but concise set of instructions dealing with the approach to the respondent and suggestions for posing certain questions and the classification of possible answers.[1]

THE QUESTIONNAIRE

THE SOCIAL SURVEY

DIPHTHERIA INQUIRY N.S.69

TOWN OR DISTRICT (as on quota)		INVESTIGATOR		DATE	
Urban	Y	Rural	X	Region	1 2 3 4
Age of Mother				*Economic Group (C.W.E.)*	
Up to 24	1	Husband at Home ..	1	Up to £3	1
25–29	2	Husband away ..	2	Over £3–£4 ..	2
30–34	3	Husband decd. ..	—	,, £4–£5 10s. ..	3
35–39	4	Divorced or Separated	3	,, £5 10s. –£10	4
40–44	5			,, £10	5
45 and over ..	6	*Last Type of Education*—		N.A.	0
		Elementary	4		
Working Full-time	7	Secondary, Technical	5	*Occupation of C.W.E.*	
Working Part-time ..	8	Others	6		
Not working ..	9				
		Substitute	1		
		Original	2		

1. Do you know what causes diphtheria?
   - Don't know 1
   - Infection from other children 2
   - Bad sanitation, dirt, etc. 3
   - Other causes 4

2. Do you know how diphtheria can be prevented?
   - Don't know 6
   - Immunisation, inoculation 7
   - Not possible to prevent it 8
   - Other ways 9

[1] The Social Survey. Diphtheria Immunisation, by K. Box. October 1945. N.S. 69.

3. Have you had your children immunised?

	Sex			Under						6	8	10	12	14		
		Boy	Girl	1	1	2	3	4	5	7	9	11	13	15	Yes	N
Selected Child	A	1	2	Y	X	0	1	2	3	4	5	6	7	8	3	4
Child	B	1	2	Y	X	0	1	2	3	4	5	6	7	8	3	4
,,	C	1	2	Y	X	0	1	2	3	4	5	6	7	8	3	4
,,	D	1	2	Y	X	0	1	2	3	4	5	6	7	8	3	4
,,	E	1	2	Y	X	0	1	2	3	4	5	6	7	8	3	4
,,	F	1	2	Y	X	0	1	2	3	4	5	6	7	8	3	4
,,	G	1	2	Y	X	0	1	2	3	4	5	6	7	8	3	4
,,	H	1	2	Y	X	0	1	2	3	4	5	6	7	8	3	4
,,	I	1	2	Y	X	0	1	2	3	4	5	6	7	8	3	4

IF SELECTED CHILD IMMUNISED
(Ask with reference to that Child)

4. How old was the Child when immunised?

Under 1 yr.	Y	3 yrs.	1	6 or 7 yrs.	4	12 or 13	7
1 yr.	X	4 yrs.	2	8 or 9 yrs.	5	14 or 15	8
2 yrs.	0	5 yrs.	3	10 or 11 yrs.	6		

5. Who suggested that the child should be immunised?

School	1
Health Visitor	2
Private Doctor	3
Welfare Clinic	4
Own idea (from publicity, etc.)	5

6. Did you have any difficulty in getting the child immunised?

Yes .. .. Y No .. .. X N.A. .. 0

If Yes, what difficulty?

7. (To everyone *not* answering Code 5 to Q.5)

How long after it was suggested, was the immunisation done?

Within a week ..	1	More than 3 months ..	4
More than 1 week to 4 weeks	2	Don't remember ..	5
More than 4 weeks to 3 months ..	3		

8. Was the immunisation completed? How many visits?

Yes .. Y No .. X

9. Was the arm painful afterwards?

Yes .. Y No .. X

10. Did you pay for it (at your own Doctor), or have it done under the Council Scheme?

Paid own Doctor .. 4 Free under scheme .. 5

IF SELECTED CHILD NOT IMMUNISED
(Ask with reference to that Child)

11. Why was the child not immunised?

Had not heard about immunisation	Y
Have not bothered, not had time, will have it done	X
Don't believe in it, not worth while	0
Husband objects	1
Bad for child, would hurt or frighten it	2
Waiting till child goes to school, not old enough yet	3
Child has just been vaccinated, is ill	4
Child has already had diphtheria	5
Waiting to consult husband	6
Tried, but unable to get it done yet	7
Other reason	8
Don't know	9

TO EVERYONE NOT ANSWERING CODE Y TO Q.11

12. Have you heard or read about immunisation in any of the following ways?

Newspapers or magazines ..	Yes ..	1	No ..	2
Radio	Yes ..	3	No ..	4
At the cinema	Yes ..	5	No ..	6
Posters	Yes ..	7	No ..	8
Leaflet or visit from school, health visitor, etc.	Yes ..	9	No ..	0

13. Do you know up to what age children should be immunised?
Right (15 yrs. old) .. 1 Wrong .. 2 Don't know .. 3

14. Do you know what is the best age to have children immunised?
Right (1 yr. old) .. 5 Wrong .. 6 Don't know .. 7

TO EVERYONE

15. Do you use a children's welfare clinic?
Regularly .. 1 Occasionally 2 Never 3 Used to .. 4

16. Have you ever been present at a school medical examination?
Yes .. .. 6 No 7

Many of the problems arising from the schedule of questions can fortunately be settled before the field work starts, and where a pilot survey is undertaken, unexpected weaknesses may be revealed. One last point is worth mentioning, only if because it is so frequently asked. Do people tell the truth when interviewed? Individuals do, of course, vary in this respect, but since anyone may refuse to be interviewed, it seems pointless for him to accept the invitation and then proceed to tell lies. In any case, consistent lying or exaggeration is extremely difficult. There is some truth, however, in the contention that not every question is truthfully answered. During the war, housewives were exhorted to salvage waste material. One inquiry which asked whether the respondent did salvage waste material revealed that a far higher proportion of housewives stated that they participated in the

campaign than the actual collections indicated. The significance of this point is that most people probably tend to state that they conform to accepted social requirements rather than admit that they do not. In a survey among housewives questions regarding pocket money and punishment of children were put. There was universal agreement with the proposition that 'children nowadays have too much pocket money' although interviewer after interviewer commented on the fact that their particular respondents never made this mistake! Similarly with the punishment of children. Apparently this was hardly ever necessary, if the answers are to be believed. As several interviewers commented: 'never were children so well behaved as today'. Clearly the respondents in many cases were giving the socially acceptable answer. It was noted that occasionally the respondent took the line, 'what business is it of yours', although up to that point in the interview she had been most co-operative. To overcome the risk of bias, conscious or otherwise, from the respondent, the schedule designer uses various devices. Thus in the I.P.A.'s Readership Survey, the following method was used to overcome the 'prestige' factor. By this is meant that respondents might claim to have read certain periodicals for prestige purposes, when in fact they never read them. It is well known, for example, that many more people 'admit' to having read *The Times* than its circulation would justify, even when allowance is made for each copy being read by several people. The following question was asked: 'When was the last time you looked at a copy of . . . . . in the past three months?' This clearly gives the respondent a loophole to admit that he has not read a particular periodical without admitting that in fact he never does read it. Only if the respondent answered this question in the affirmative were further questions as to place and time asked of him. Much the same reasons explain the phrasing of the following question 'Do you have a T.V. set at home YET?' rather than the same question without the final word. The inference is that the respondent can afford it, but has not decided so far to buy a T.V. set.

The Problem of Interviewing

It has been stated that the 'representativeness of a sample depends on the ability of field workers to trace their subjects and

persuade them to co-operate in the completion of a questionnaire; and the accuracy of the results depends on that of the information recorded. Much hinges on the address, skill and tact of the interviewer, who thus becomes a possible source of serious bias in the enquiry.'

The majority of sample surveys are conducted by interviewers. A good interviewer can persuade his subject to reply to almost any question and it is fortunate, if somewhat surprising, that there appears to be no limit to the variety of questions which the average person is prepared to answer. Such willing co-operation can only be attributed to a general human weakness of being flattered by others seeking one's views. Occasionally, the respondent may be interested in the subject of the survey and be especially willing to participate, particularly if he feels that his opinions may influence the attitude of others. The experience of the Social Survey, some of whose interviews may take up to an hour, is that on an average only 3 per cent of the original sample members refuse to be interviewed. Unfortunately, all interviewers are not of equal ability, although most of the survey bodies now employ only trained personnel. The proportion will, of course, vary from enquiry to enquiry, dependent on what is expected of the respondent, but it is seldom large. One exception was the high refusal rate – approx. 60% of the respondents contacted – in the National Food Survey.[1] The reason for this lay in the fact that the housewife was expected to keep a detailed account of food purchases for a week and allow the interviewer to check the contents of her larder.

Just how significant training of interviewers may be in affecting the response of informants and the quality of that response, is difficult to judge. The results of a detailed investigation to test the relative efficiencies of two groups of professional investigators from the Social Survey and the British Institute of Public Opinion on the one hand and a group of University students on the other suggest that training may not be all that important.[2] It appears from this enquiry that while the professionals enjoyed a higher success rate in establishing contact with the respondents than the students, the relative differences between the three groups of interviewers with regard to the quality of response in

[1] Domestic Food Consumption and Expenditure, 1950. H.M.S.O. 1952.

[2] J. Durbin and A. Stuart 'Differences in Response Rates of Experienced and Inexperienced Interviewers'. J.R.S.S. 1951, Part II.

terms of completed and accurate schedules were not such as to warrant the inference that the students were much inferior to the professionals. There was some evidence, however, to suggest that on the more difficult questions the students were not as effective as the professionals. In view of the present trend towards longer schedules and intensive interviewing of the respondent, *i.e.* 'depth' interviewing to seek out causes and reasons for attitudes, the importance of training will become more evident. In the early stages of survey work, the main quality required of an interviewer was 'personality' in the sense that she could easily establish 'rapport' with her informant and persuade him or her to talk freely on the survey topic. Nowadays much more attention is being paid to sequence of questions, the form in which they are posed, and more skill and concentration is required of the interviewer.[1]

Apart from the difficulties already discussed in connection with the questionnaire, the actual interview is attended by even greater problems. The simplest is the risk that the interviewer may misunderstand a reply, or merely mark off the wrong code number for any pre-coded answer. The risk of misinterpretation is greater with opinion questions, where the main replies have not been classified in advance on the schedule, than with questions of fact. More important, however, is the actual conduct of the interview itself. There is always the danger of prompting the hesitant respondent at the wrong time, even putting the answer to him. A particular cause of concern is the extent of what is termed 'interviewer bias.' 'Bias', in the normal sense of the word, is a more serious danger with the public opinion polls than with social or market surveys. The Princeton Office of Opinion Research has carried out many tests in this connection and has revealed that even with professional interviewers bias is unavoidable. In one American election, professional interviewers were divided into groups of opposite political faiths. Their returns revealed quite clearly the effects of their subconscious sympathies on their respondents. Nor is this a new problem; it has been known to exist from the earliest days of sample surveys. For example, as long ago as 1914 an American sociologist found that the replies of 2,000 destitute men explaining their distress were markedly influenced by the interviewers' sympathies and

[1] Fothergill and Willcock, *op. cit.*

views. Thus a Prohibitionist interviewer's results revealed a strong tendency among his respondents to attribute their destitution to drink; an interviewer with Socialist leanings recorded many who ascribed their position to industrial causes. According to the author, 'quantitative measures of interviewer bias in this particular survey turned out to be amazingly large. The men may have been glad to please anyone that showed an interest in them'.[1] The same authority comments that interviewers will influence their respondents' replies by the mood into which the latter are put. For example, 'the interviewer may make the respondent gay or despairing, garrulous or clammish. Some interviewers unconsciously cause respondents to take sides with them, some against them'. As Mr Deming points out, training can do much to overcome the more obvious causes of interviewer bias, but even with a well trained corps of interviewers it may arise, even quite unconsciously. The experience of the Social Survey in the course of collecting data on the probable response of ex-Service men to the offer of 1939–45 Campaign Stars is noteworthy in this respect. It was found that the age of the female interviewer had a definite influence on the attitude of the male respondent. The younger the interviewer, the more likely was the man to disavow any intention of applying for the awards!

Despite the intensive training given to interviewers, the holding of briefing conferences and the issue of detailed instructions with the schedule setting out the considerations which have prompted the various questions and the best way of putting them to the respondent, mistakes are inevitable. It is not the obvious mistakes and glaring inconsistencies between answers that are troublesome. These can usually be detected in the editing of the schedule. It is the minor slips, such as incorrect interpretations of an answer or poor classification of the respondent which are so difficult to eliminate, since they cannot be detected and the results are thereby distorted. When all the problems involved in the organisation and conduct of a survey are taken into account, the errors occurring at the interview remain the most serious. In the words of two members of the Social Survey, 'sampling errors are the least serious, it is the human errors such as errors in classification and memory errors on the

[1] 'On Errors in Surveys', W. E. Deming in *American Sociological Review*, August 1944.

part of the respondent . . . that are less easily detected'.[1] A more recent enquiry into this problem carried out by the Social Survey revealed that over three-quarters of the mistakes made by investigators during the course of the interviews (and noted by observers present at the interview) could not have been detected from a scrutiny of their schedules when returned to the head office.[2]

Some idea of the task imposed upon the investigator is given by the form below headed 'Classification'. This is actually a supplementary questionnaire to the main questionnaires which were concerned with surveys of the public's knowledge of and attitude towards tuberculosis, on reading habits and on savings. The purpose of this supplementary list of questions was to provide information concerning the informant's living conditions, social class, household composition etc. independently of the three main surveys. The questions are quite clear and it will be noted how the investigator seeks to ascertain the respondent's income group. The respondent is not actually asked what he earns, but is asked to indicate to which particular income group, as given in Question VIII, he belongs. The letters S.W.E. stand for senior wage earner, usually the male head of the household. Since it may not be possible to interview him if the investigator calls during working hours, the 'subject', usually his wife, will provide the information. The interviewer will need briefing on the appropriate method of answering Question X; assessments by the individual directly concerned of his success or otherwise are not usually very satisfactory for comparative purposes, *e.g.*

CLASSIFICATION † *

(i) Interviewer's name ....

Interviewer's number . . . . / . . . /. .

(ii) RING DATE OF INTERVIEW.

	Sun.	Mon.	Tues.	Wed.	Thur.	Fri.	Sat.
April	16	17	18	19	20	21	22
	23	24	25	26	27	28	29
	30	—	—	—	—	—	—
May	1	2	3	4	5	5	6

(iii) Subject. Where living.

At home	Y
In institution, hotel	X
As a boarder	0
In rooms	1
As resident servant	2

(iv) Type of dwelling.

Detached house	3
Semi-detached house	4
Terraced house	5
Self-contained flat	6
Part of house	7
Other (*specify*)	8

....

[1] Gray and Corlett, *Sampling for the Social Survey*, J.R.S.S., 1948, Part II.

[2] Fothergill and Willcock, *op. cit.*

(v) HOUSEHOLD COMPOSITION

Relationship to Subject	Age	Sex M	Sex F	Status M	Status S	Status W	Working F	Working P	Working N
A. Subject		1	2	3	4	5	6	7	8
B.		1	2	3	4	5	6	7	8
C.		1	2	3	4	5	6	7	8
D.		1	2	3	4	5	6	7	8
E.		1	2	3	4	5	6	7	8
F.		1	2	3	4	5	6	7	8
G.		1	2	3	4	5	6	7	8
H.		1	2	3	4	5	6	7	8
I.		1	2	3	4	5	6	7	8
J.		1	2	3	4	5	6	7	8
K.		1	2	3	4	5	6	7	8

Total 0—4 ....

Total 5—15 ....

Total 16 and over.. ....

Total in household...........

Housewife is (give letter) ....

S.W.E. is (give letter) ....

(vi) Number of rooms . ...

(vii) Subject occupation (full description)

....

....

Subject industry, trade or profession ...

....

Self employed 1
Employee 2

S.W.E. occupation (full description)

....

....

S.W.E. industry, trade or profession .....

... .

Self employed 1
Employee 2

(viii) Income per week less deductions plus bonuses.

		Sjt.	S.W.E.
	Nil	0	0
	Up to £3	1	1
	Over £3 to £5	2	2
	Over £5 to £7 10s.	3	4
SHOW	Over £7 10s. to £10 ...	4	4
CARD	Over £10 to £20	5	5
	Over £20	6	6
	Don't know	7	7
	Refusal, not asked . . .	8	8

If SJT. D.K., REFUSAL, NOT ASKED
Why? .. . .

.. ..

If S.W.E. D.K., REFUSAL, NOT ASKED
Why ? ... .

....

(ix) Interview situation.

CODE ALL THAT APPLY

Informant alone ....
Spouse present 1
Other adult(s) present 2
Children present 3

(x) Interviewer's assessment of success of interview.

Above average (give reason) 4
Average 5
Below average (give reason) 6
Very poor (give reason) 7

....

....

(xi) Serial number on record sheet...............

(xii) Subject education.

Age left school ....
Type of last school.
Elementary ....Y
Central, Technical, Commercial...X
Secondary, Public 0
University........... 1

† Reproduced by permission of Director of Research Techniques, London School of Economics and the Editor of the Journal of the Royal Statistical Society.

* *Source:* Durbin and Stuart, *op. cit.* This particular paper contains three schedules of questions which will repay study by the student. The discussion of the paper is also useful.

comparing the performance of interviewers. The question does give some indication of the tenor of the interview and indicates the degree of co-operation which the investigator received from the informant. In passing, it should be noted that the interviewer is also required to keep a list or record of all his calls, whether or not successful. When one takes into account the physical strain of making contact with respondents in all weathers and at all times, as well as the nervous strain of continuous interviewing, it is not surprising that mistakes are made by investigators.

Postal Enquiries

Quite apart from the problems and weaknesses of personal interviewing as a survey method, an important consideration is cost. A cheaper method is to use a postal questionnaire. At first sight the postal questionnaire has several advantages. It can be sent to a very large number of people at low cost, so that the sample size may be increased considerably, relatively to the survey sample for the same cost. Further, the risk of bias or mistakes on the part of the interviewer is absent. All these apparent advantages, however, prove on examination to be fictitious. The fundamental weakness of the postal method is the low proportion of returns. On the average, a 20 per cent response is considered good. It may be argued that if 100,000 schedules can be sent out for the same cost as interviewing 5,000 people, the final sample still contains 20,000 and is therefore preferable. Unfortunately, there is no means of ascertaining whether or not the 20 per cent return constitutes a representative sample of those people to whom the schedule was sent. It almost certainly does not. It is probably true to say that the cost per completed schedule by interviewing is ultimately little above that for the postal enquiry. There is, too, the greater reliability of the former, since despite the risk of interviewer bias and mistakes, the respondents will themselves make mistakes in completing the forms.

The experience of the Social Survey with postal schedules is worthy of note. Some 16,000 members of a profession were circularised and 38 per cent returned the forms immediately. A further 32 per cent replied, after a reminder had been sent. Ultimately, the response was just over 80 per cent. Such an experience was quite exceptional for the following reasons. The subject

of the enquiry was connected with the future of the profession, a matter of considerable interest to the members. They would in any case return a higher proportion than would be received in an enquiry on any subject covering the general population for, as professional people, their reaction to forms would not be that of the average member of the public. Finally, a pilot survey had provided many useful lessons and the schedule of questions was devised with very great care. The last points should not be exceptional, but the fact that the organisers of the survey comment on them suggests that they are not always accorded the same degree of attention.

Another enquiry undertaken by the same body followed the more usual pattern. Of 740 households sampled and circularised, only 16 per cent replied until a first reminder produced replies from another 21 per cent. After a second reminder, 8 per cent more replied, making up a total response of 45 per cent after two reminders. Even this was very satisfactory, but here, too, the householders could be assumed to have a personal interest in the subject of the enquiry. In the light of these examples the probable response to a commercial questionnaire requesting opinions on, for example, a manufacturer's soap, may be imagined!

The obvious method of overcoming this weakness of postal surveys is to make the returns compulsory. The only body with statutory powers to compel returns is the Government. In consequence, they are the main users of the postal enquiry, and use it for the censuses of population, electors, production, distribution and earnings. It should not be assumed, however, that because the recipient is compelled to make a return, he necessarily completes the schedule accurately. The census of distribution forms carry a note to the effect that if exact figures are not available, then an estimate may be inserted. Such minor weaknesses are of little significance since small errors will tend to cancel each other out in the aggregate. This assumes, however, that *only* minor slips are made. It is doubtful if this is true when the schedules are virtually small booklets, as used to be the case in the censuses of distribution and production, completion of which is all too often regarded as an imposition to be dealt with in the minimum space of time and with even less care. A more important consideration is that the type of information required by

the Government Departments is purely factual. Questions which can be answered by either 'yes' or 'no', or a figure can be posed equally well by post as by interviewer. But the same is not true of opinion questions, consequently most surveys seeking other than bare factual statements must be carried out by interviewers. It is noteworthy that the Verdon Smith Committee in its 1954 report (Cmd. 9276) on the Censuses of Production and Distribution made certain recommendations in this respect which have been adopted by the government. In 1955-7 inclusive there were 3 sample surveys on Production covering only 35,000 firms instead of about 250,000 in a census. Furthermore, the questionnaire has been reduced to a single sheet containing some twenty questions similar to that reproduced on p. 33. Such changes are a direct consequence of the criticisms directed at the census authorities by industry and trade.

Conclusions

The foregoing sections are designed to provide a summary outline treatment of the conduct of a sample survey. Although the student's attention has been directed to one aspect and problem of survey design at a time, he should never forget that each survey should be regarded as an integrated whole. Despite all the emphasis in the preceding pages given to sample and questionnaire design as well as interviewer bias, the main problem confronting the survey organiser is first, to determine the nature of the problem, and secondly to decide how, for a given outlay, it may be answered most effectively. The early chapters in this book emphasised the need to verify the quality and nature of any statistical data irrespective of its source. This warning applies with especial force to the results of economic and sociological surveys. The existence of a few figures seems to encourage statements the validity of which rests entirely on such limited evidence. Data derived from such surveys should not be used unless the report has itself been scrutinized. Any reputable survey body will publish in the report of its findings some account of the method employed to select the sample, and the classification of the respondents which was employed. A copy of the questionnaire should be given as well as the instructions to the interviewers.

In brief, all the difficulties cited in the preceding paragraphs

should find some mention.[1] The reader may then be in a position to estimate just how satisfactory the data are and make the best use of them.

[1] The following Reports, to some of which references have already been made, are all good examples in these respects. The explanatory manual of the National Readership Survey 1957, I.P.A. is particularly interesting since it gives a full account of the sample design, the interviewing technique as well as a copy of the schedule. There is a description of a survey undertaken by the Social Survey in Appendix A of the Report on Domestic Food Consumption and Expenditure 1953, H.M.S.O. 1955. The Household Expenditure Enquiry 1953/4 report also has a clear description of the methods used in that survey. A detailed account of all the problems encountered in the fertility enquiry prepared by Glass and Grebenik on behalf of the Royal Commission on Population entitled 'The Trend and Pattern of Fertility in Great Britain', Part I, illustrates many of the practical difficulties encountered both in the planning and the interviewing stages. One of the best accounts of a survey, which discusses not merely the practical problems, but at the same time explains many of the theoretical problems of sample design, is given in *British Income and Savings* by H. Lydall, Chapter 7. Particular attention should be paid to the schedule and the problem of the high non-response rate in this enquiry. The student will find the account extremely rewarding

CHAPTER XIV

SOCIAL STATISTICS

The administrative needs of the government extend beyond the fields of industry and economic planning. A major part of the government machine is concerned with the provision of social services and, as with economic statistics, the administrative processes yield a large crop of statistical data. Most of these are to be found in the annual reports of the appropriate Ministry, *e.g.*, the two-volume report of the Ministry of Health or that of the National Assistance Board. As with economic data, some enquiries are undertaken for special purposes. For example, the annual survey of food consumption conducted by the National Food Survey, the enquiry into household budgets undertaken by the Ministry of Labour in 1953-4, as well as the survey into 'Reasons for Retirement' among the retired population, conducted by the same Ministry.[1] A wealth of data relating to social conditions is to be found in the Census of Population reports. Some of the methods used for analysing vital statistics are discussed in the next chapter.

This chapter is designed to bring to the reader's attention some of the better known and more important social statistics, many of which are frequently quoted in the Press. Inevitably in a book of this type, the coverage must be limited. Fortunately for the interested reader there is an authoritative up-to-date survey of social conditions based entirely upon the relevant statistical data. Every student of the social sciences should read it.[2] A word of warning is, however, necessary. While tables illustrating the form of the published data are reproduced in this chapter and in the aforementioned book, it is quite impossible for the reader to obtain more than a mere notion of what statistical information these government reports offer. If the reader requires more than the brief outline offered in the pages which follow, he should consult these reports at his local reference library. If

[1] Published H.M. Stationery Office, 1957.

[2] *A Survey of Social Conditions in England and Wales as Illustrated by Statistics*, by Carr-Saunders, Caradog Jones and Moser, 1958.

possible he should try to extract the answer to a particular question on the subject; *e.g.*, how many people drawing the National Insurance pension applied for National Assistance supplementation, or what is the average working week and hourly wage in the shipbuilding industry? Finding the answer to small but precisely defined questions such as these provides the opportunity to study with just that degree more care than a mere desultory flicking-over of pages and is infinitely more illuminating in revealing the deficiencies of the published data.

Housing

There would be little disagreement with the assertion that the major social problem in post-war Britain has been the shortage of housing. For this reason it is rather remarkable, to put it no more strongly, that the available statistics are totally inadequate to form the basis of a rational housing policy. Quite apart from the complications arising from the system of rent control which made a true assessment of the adequacy or otherwise of the existing housing accommodation virtually impossible, the statistical shortcomings are very obvious. There is first no precise data relating to the supply and adequacy of existing housing accommodation, *e.g.*, age and amenities; second, the statistics to measure either 'need' or 'demand' are inadequate.

The data relating to housing are derived from two main sources. The first is the Ministry of Housing which prepares data relating to the volume of current building. The second is the Census of Population which, in the Housing volume, provides a great deal of interesting information relating to housing conditions and over-crowding. The two sets of data are best dealt with separately. The statistics of house building date back to the end of the first World War when in 1919 the government first made itself responsible for the provision of subsidised dwellings. Before then, the number of dwellings could only be estimated from census returns. Once public money was at stake, the administration annually produced statements showing the number of houses built which were entitled to the subsidies. From 1922 on, the published data became virtually comprehensive since it included all non-subsidised accommodation with a rateable value of £78 and under.[1] With the tremendous boom in house

[1] In the Metropolitan Police District the limit was £105 but houses with Rateable values above £78 in the provinces and over £105 in the M.P.D. were very few.

construction of the 1930's the Ministry of Health (which from 1919 was responsible for this matter) published as from September, 1934, a half-yearly return giving the number of houses built by both private enterprise and local authorities, as well as the volume of 'slum-clearance'. A particularly useful and informative publication of this period is the 'Survey of Over-crowding' of 1936 which covered Scotland, as well as England and Wales.

The end of the second World War witnessed a revival of government housing activity but on a much larger scale than in 1919. Between 1946-48 a monthly return of house construction and war damage repair in Great Britain was produced by the Ministry of Health. Thenceafter it was converted into quarterly returns for England and Wales (separate returns for Scotland) analysing the number of houses built by reference to the region, and the authority responsible for the construction, *e.g.* local authority or private builder, etc. From 1951 these *Housing Returns* were prepared by the new Ministry of Housing and Local Government. In addition, these returns included data relating to the number of tenders approved, construction begun, and completions. In addition to the data on permanent houses, in recent years similar data are provided relating to conversions and adaptations of old and existing property, as well as the progress in slum-clearance.

The fact that statistics of housing construction have been officially prepared since 1919 should not delude the reader into thinking that the data form a consecutive series which reflect the pattern and fluctuations in house building over four decades. Miss Marion Bowley has emphasised in her study of housing statistics that these data abound with snags such as changing definitions, varying coverage of series, etc., and therefore require the greatest care in extraction if any period comparisons are required.[1] For all practical purposes the various Housing Returns serve as an indication of the scale of housing construction at various periods in England and Wales and little else.[2]

The second source of housing statistics is the Census of Population. Table 33, illustrates for successive censuses some of the available data. The Housing volume of the Census of Population 1951, is the main current source, but details relating to

[1] Housing Statistics of Great Britain, *J.R.S.S.*, 1950, Part III.

[2] The most detailed official account of housing statistics, *i.e.*, building, is given in an article in *Economic Trends*, June 1958.

housing conditions are given in the County volumes as well as in the 1 per cent Sample Tables. The 1951 Census was especially important in this particular respect for it collected new information relating to what are called 'household arrangements', *i.e.*, the existence of piped water supply, cooking stove, water-closet and fixed bath, as well as providing more detailed and new classifications of households not given in the 1931 volume. For example, households are analysed by the number of earners and number of children they contain, as well as the proportion with the 'household arrangements' specified above.

TABLE 33

POPULATION, HABITATIONS AND HOUSEHOLDS, 1861–1951* ENGLAND AND WALES

Census Year	Population 000's	Houses 000's	Families 000's	Persons per House
1861	20,066	3,924	4,492	5·1
1871	22,712	4,520	5,049	5·02
1881	25·974	5,218	5,633	4·98
1891	29,003	5,824	6,131	4·98
1901	32,529	6,710	7,037	4·85
	Population in private families 000's	Private dwellings 000's	Private families or households 000's	Persons in private families per dwelling
1911†	34,606	7,691	7,943	4·50
1921	36,180	7,979	8,739	4·53
1931	38,042	9,400	10,233	4·05
1951	41,840	12·389	13·118	3·38

* *Source: Economic Trends.* June 1958.

† In the 1911 and subsequent censuses non-private households were excluded from the tabulations which in previous censuses had included them.

The term 'housing' in the Census volumes does not merely cover houses as such; it is primarily concerned with available accommodation, the basic unit of which is defined as a 'structurally separate dwelling'. This has been defined since the 1921 census as follows: 'any set of rooms, intended or used for habitation, having access either to the street or to a common landing or staircase'. According to the 1951 Census there were 13·3 million 'structurally separate dwellings'. The feature of the census data is not so much the actual enumeration of such accommodation but the extent to which it is occupied. In other words, the term 'housing' as used in the Census would be more

accurately defined as 'housing conditions'. The 13·3m. structurally separate dwellings were occupied in 1951 by 14·5m. households; of this number 2·1m. were sharing 0·9m. dwellings.

The term 'household' is also carefully defined in the Census. To start with they are divided into two main groupings, private and non-private. The former consist of what is generally known as a household, *i.e.*, any group of individuals such as a family and servants and lodger with board. The non-private households covers all the people living in institutions and hotels, etc. In 1951 this latter group comprised less than 5 per cent of the population, the balance living in private households. Of their number the Census provides the following analysis by size:

No. of persons in household	1	2	3	4	5 and more;
Percentage of households	11	28	25	19	17

and the comment is added that compared with 1921, the proportion of households with 1–3 persons had risen from 45 to 64 per cent.

The private households are in their turn further sub-divided into 'primary family units' and what are termed 'composite' households. The former consists of the head of the household, his wife, their children (of any age if unmarried) and any immediate relatives such as brothers and sisters of the head of the household and his wife. Resident domestic servants are also included. In other words, the primary family unit is what most of us regard as the 'family'. The 'composite' household consists of a private family unit plus one or more of the following, a family nucleus; a married brother or child of the head of the household; and any others not related to the head of the household, *e.g.*, a boarder. The term 'family nucleus' signifies any group in a composite household such as a married couple, with or without children, or a lone parent, *e.g.*, widowed or divorced. The importance of the concept 'family nucleus' arises from the need to make some estimate of housing needs. Such 'nuclei' can be regarded as potential occupiers of their own establishment; they may only be living with the primary family unit until such time as they can acquire separate accommodation. About four-fifths of such nuclei were in fact married sons or daughters of the head of the households, awaiting suitable accommodation for themselves.

An important feature of the Census data is the information it provides on the subject of 'over-crowding', or what the Census defines as 'density of occupation'. 'Overcrowded' is a purely subjective term, and for statistical purposes it is essential to have some standard, although such a standard is quite arbitrary.[1] The standard which has been used in the 1931 and 1951 Census reports provides for a maximum of two persons per 'habitable' room.[2] A density greater than this figure qualifies for the description over-crowded. The very great increase in house building since 1931 has much reduced this problem. In 1951 just over 1 per cent of the population in England and Wales was living in overcrowded conditions compared with a percentage over three times as great in 1931. The Census data provide a regional breakdown of the global figures showing in which regions and towns the problem is most serious. Directly related to the problem of over-crowding (in terms of number of persons per room) is the proportion of shared dwellings. As stated earlier, just over 2 million households share their accommodation with other families; most of those households were small, *i.e.*, 1-3 persons. According to the General Report over 1½m. of these were not 'seriously deficient' in numbers of rooms in relation to the households contained in them.

Admittedly the term 'overcrowding' is arbitrarily defined in any enquiry, but there can be little doubt that the Census definition is out of date and under-estimates the extent of the problem in so far as it sets a standard which is unduly low. Since 'room' covers also the kitchen if used as a living room (the Census does not distinguish between bedrooms and living rooms, nor are their dimensions measured), then a three-bedroomed house with sitting-room and kitchen-living-room would not be overcrowded until it had over ten persons. No attempt is made to take the sex or age composition of the household into account in this respect. A better criterion for over-crowding would be to use a 'bedroom' standard. This would aim to provide adequate sleeping accommodation to allow separation of the sexes over say 10 years of age (except of course for married couples) and smaller sleeping accommodation for young children, *e.g.*, one room for

[1] Compare for example the detailed standard used in the 'New Survey of London Life and Labour 1928' with the standard employed in the 1936 Overcrowding Survey.

[2] The Census does not define a habitable room in precise terms, except to exclude sculleries and bathrooms.

3 children under 5, or one child under with one over 5, or two children aged 5 or over of the same sex. This implies that a working kitchen and separate living room are essential during the day and should not be occupied at night. An even more precise standard could be calculated by designating each adult, for example, as a unit and adolescents as 3/4, children under 10 as half and infants under 1 as 1/4 or nil units, and relating the household total to the floor space available. Then a given-sized house should not contain more than so many units per x square feet. Another problem is of course the size of the rooms; generally speaking the poorer housed the household by the foregoing standards, the smaller are the actual 'rooms'. Any survey of over-crowding or 'housing conditions' should be examined to ascertain what standards are used. For example, the Housing Act 1936 used both a 'bedroom' and 'floor area' standard to define overcrowding. This is deemed to exist when the number of persons sleeping in the house is *either* such that two or more of those persons being over ten years old, of opposite sex and not living together as man and wife, must sleep in the same room; *or*, that there is an excess of the permitted number of persons as ascertained in relation to the number and floor area of the rooms as laid down in a schedule to the Act, For this purpose an infant under one year is not counted and a child between 1 and 10 years is reckoned as half a unit.

The collection of information in the 1951 Census relating to 'household arrangements' revealed a far from satisfactory state of affairs. Just over half of all households in England and Wales have all these four amenities and exclusive use thereof. There were 1·4m. households without exclusive use of *both* a kitchen sink and a water closet. As is to be expected, large numbers of households in shared dwellings were without separate provision of these domestic arrangements. Even so, over one third of private households had no fixed bath, 8 per cent had no water closet, 6 per cent had no kitchen sink and the same proportion was without piped water. Note that Table 34 gives the percentages of households without *exclusive* use; not as in the preceding sentence, *entirely* without.

A study of the Housing volume or even the Housing section in the 1 per cent Sample Tables alone, will reveal many detailed analyses of households by reference to their composition by

TABLE 34

REGIONAL DISTRIBUTION OF OVERCROWDING AND HOUSEHOLD ARRANGEMENTS*

	% of persons living more than 2 per room		Persons per room		Persons per household		1951 only % of households without exclusive use of:		
	1931	1951	1931	1951	1931	1951	W.C.	fixed bath	stove and sink
England and Wales	6·94	2·16	0·83	0·73	3·72	3·19	21	45	14
Scotland	n.a.	15·47	1·27	1·04	3·99	3·39	35	49	15
Northern†	16·90	4·54	1·02	0·84	3·97	3·37	19	45	17
Midland†	7·21	2·80	0·85	0·76	3·92	3·36	23	46	15
Southern	2·78	1·59	0·71	0·70	3·60	3·21	22	40	13

* *Source: 1% Sample Tables. Table IV.7.*

† Inter-year comparisons affected slightly by boundary changes.

number, sex and age. Table 34 above indicates the type of data published. The student is advised to examine these tables in the Census volumes so that the nature of their contents at least can be impressed upon his memory. No second-hand summary such as given above can replace such a scrutiny in bringing home to the reader the coverage of such tables.

REFERENCES

Census 1951. 1 per cent Sample Table, Part 1.
Census 1951. Housing Report, Ch. VI.

Social Class

The English, it has been observed, are extremely class conscious. Certainly most people are ready to classify others by reference to their social class, such as working, middle or upper class. There is, nevertheless, no single measure of 'class'; for example, different people might well classify a particular individual either as working or 'lower' middle class. The conventional basis for social classification is well understood, but it is not sufficiently precise for statistical purposes. In any case many people dislike 'class consciousness' as such, and would reasonably ask why it interests the social statistician. There are, however, a number of reasons why the subject is important. First of all, it is common knowledge that fertility among labourers is higher than among the professional community; just

as is mortality. It is not sufficient merely to believe that a difference exists between the two extreme groups. We want more precise information as to its extent. For example, before the war, in certain districts of England and Wales the infant mortality was almost four times as high as in the wealthy areas of the South East.[1] This is criminal waste and once we have the information as to the extent thereof, social policy can be adapted to cure such blots on the community. It is a commonplace to talk of the opportunity open to the poor boy to get 'to the top'. But how many get there? To what extent is there what the sociologist calls 'social mobility', *i.e.*, do sons and daughters generally move into a higher or lower social class than their parents?[2] Before such enquiries are practicable, there must be a generally accepted classification of social status.

There are, of course, a number of criteria which are employed to classify people. Accent is often a good guide; but statistically it is useless. Education is a useful indicator but is not enough. Income is a better guide and lends itself to statistical treatment; but there are many people with a great deal of money these days whom no one would place in the 'upper class', just as few people would classify a parson in a poor living as 'working class'. The most generally acceptable basis for classifying people is their occupation. Furthermore, surveys have shown a high degree of unanimity among all classes as to the relative status of given trades, occupations and professions.[3] Since the 1911 Census the Registrar General has used a five-fold classification as follows:

I. Professional, etc., occupations, *e.g.*, doctor, lawyer.
II. Intermediate occupations, *e.g.*, business executive, manager of large store.
III. Skilled occupations, *e.g.*, draughtsman and policeman.
IV. Partly skilled occupations, *e.g.*, ticket collector and plumber's labourer.
V. Unskilled occupations, *e.g.*, dock labourer, watchman.

Such a classification is a good deal better than the rough and ready three-fold classification, working, middle and upper! It is precise, but for that reason it is arbitrary. For most people the

[1] See *Population and Poverty*, R. Titmuss, for a study of pre-war society based entirely on such data. Even now there are regional differences, see for example Table 36 below.

[2] See *Social Mobility in Britain*, ed. D. V. Glass

[3] *Social Mobility in Britain*, *op. cit.*

members of classes II and III inevitably merge one into the other, and in fact the latter is so large that it is not really homogeneous, *i.e.*, it is too mixed. For this reason in the 1951 Census publications the Registrar General introduced a new and more detailed classification by 'socio-economic' groups. There are 13 mutually exclusive groups within three main groupings, agriculture, non-manual, manual, and a single group for the armed forces (see Table 35). Like the social class classification, the socio-economic grouping is based upon occupations, but this more detailed classification ensures a greater degree of homogeneity within the individual group.

TABLE 35

STANDARDISED MORTALITY RATIOS FOR MEN, MARRIED WOMEN AND SINGLE WOMEN AT AGES 20-64 BY SOCIO-ECONOMIC GROUPINGS*

Socio-Economic Group†	Males	Married Women	Single Women
Agriculture			
1. Farmers	70	93	72
2. Agricultural workers	75	95	64
Non-Manual			
3. Higher administrative etc.	98	96	82
4. Other administrative etc.	84	81	70
5. Shopkeepers	100	99	97
6. Clerical workers	109	91	75
7. Shop assistants	84	79	82
8. Personal service	113	101	84
Manual			
9. Foremen	84	91	86
10. Skilled workers	102	105	109
11. Semi-skilled workers	97	108	99
12. Unskilled workers	118	111	103
All occupied and retired	100	100	85

* Group 13 of this classification – the Forces – is omitted.

† *Source: Reg. Gen. Decennial Supplement 1951. Occupational Mortality, Part II, Vol. I.*

Tables 35 and 36 illustrate the use to which the Registrar General puts these classifications. The former table, which also serves to indicate the socio-economic grouping referred to above, reveals the difference in mortality for three sections of the community at ages between 20 and 64, *i.e.*, males, married and single women. The figures themselves are known as Standard Mortality Ratios. These are explained in detail in the next chapter, but they are in effect indices based upon the relative mortality experience

of each group, allowance having been made for age differences between the groups. Farmers have the lowest mortality among males and those engaged in personal service the highest. The married women, *i.e.*, wives, have a broadly similar experience to their husbands, but note that farmers' wives do not enjoy the same favourable position of the males. In the data for single women, the figure for clerical workers is most marked in relation to other two groups, *i.e.*, men and married women.

Table 36 gives a breakdown by region and class of the infant mortality rate. Note the two rates; one relating to deaths under four weeks known as *neo-natal* mortality and the other to deaths between four weeks and one year referred to as *post-natal* deaths. The division is significant since with the marked fall in the infant mortality rates in recent decades, the main hope for its further reduction is to be found in reducing the *neo-natal* rate. The conclusions to be drawn from this table are quite clear. The regional differences are obvious, but are nothing like so marked as the inter-class differences. These tabulations of mortality by social class are extended to individual industrial groupings for certain major diseases. For example, cancer rates for different ages within individual occupations are calculated, as well as for many other causes of death. The reader should consult the relevant volume, 'Occupational Mortality', for details.

TABLE 36

NEONATAL AND POST-NEONATAL MORTALITY RATES PER 1,000 LIVE BIRTHS BY SOCIAL CLASS IN FOUR REGIONAL GROUPS

	Social Class					
	I	II	III	IV	V	All Classes
Aged under 4 weeks:						
North	12·5	16·9	19·7	21·8	23·5	20·2
Midland and East	14·9	15·6	18·0	19·8	22·8	18·5
South	12·0	15·9	15 9	18·7	19·2	16·5
Wales	16·2	21·3	19·1	24·7	26·9	21 6
Aged 4 weeks and under 12 months:						
North	4·9	7·6	13·3	18·1	23·2	14·8
Midland and East	5·2	6·5	10·3	12·8	18·1	10·8
South	5·0	4·6	7·8	9·3	13·2	8·0
Wales	4·3	7·3	13·4	16·9	22·2	14·1

Source: Reg. Gen. Decennial Supplement 1951. Occupational Mortality, Part I.

The Registrar General pays just as much attention to social class differences in fertility as in mortality. In some respects the former is more interesting to the sociologist. There have been several enquiries into fertility, in particular the questions asked in the 1911 and 1951 censuses, and the Family Census taken in 1946 for the benefit of the Royal Commission on Population. The latter is published as Volume VI of the papers of the Royal Commission under the title 'The Trend and Pattern of Fertility in Great Britain'. Table 37 is taken from that report. It shows the average number of children borne by a woman marrying between the ages of 20 and 24 in the quinquennium 1920-24. The data tell their own story; as we go down the social scale so fertility increases with the exception of the salaried employees and non-manual wage earners. This is probably due to the fact that these two groups come into close contact with more fortunate economic groups and have a greater interest in restricting the size of their families in order to maintain their living standards. Note the slightly different classification based on nine occupational groups used in this enquiry. This was devised before the Registrar General's 1951 socio-economic classification to give a more detailed breakdown of social classes than the social class classification used in the 1911 census. The 1951 Census volume

TABLE 37

GREAT BRITAIN: FAMILY SIZE FOR WOMEN MARRYING IN 1920-4 AT 20-24 YEARS OF AGE*

Status Category	Number of live births per woman
Professions	2·02
Employers	2·13
Own account	2·28
Salaried Employees	1·90
Farmers and Farm Managers	2·78
Non-manual wage earners	2·26
Manual wage earners	2·96
Agricultural workers	3·14
Labourers	3·76

* *Source: The Trend and Pattern of Fertility in G.B.*

on Fertility has not yet been published, but the 1 per cent Sample Tables give an analysis of fertility according to the socio-economic groupings. These reveal that the clerical workers' 'index' of fertility (actually a standardised ratio) was

only 84 in 1951 compared with 124 for unskilled labourers, while the higher administrative, professional and managerial group had an index of 90.[1]

Generally speaking, the classification of respondents in most sample surveys is based upon income, which is closely correlated with social status. In quota sampling, an occupational classification is generally used (see pp. 192-4). Even the Registrar General has used a simple classification by salary/wage earner, but this was done primarily to help the national income statisticians. For official studies the Registrar General's classification is employed although a special classification may be devised for a particular survey, as in the Glass study of social mobility. But the differences were negligible. Apart from the publications to which reference has been made the reader wishing to pursue his studies in this field further may find the references below of some interest. They illustrate the uses to which such data can be put in describing the social scene.

REFERENCES

The Changing Social Structure of England and Wales 1871-1951. D. C. Marsh.
Studies in Class Structure. G. D. H. Cole.
Articles on Social class in the *British Journal of Sociology* during 1957-8.

Crime

The main source of statistics relating to crime is the annual report entitled 'Criminal Statistics' prepared by the Home Office. Separate volumes are prepared for England and Wales and for Scotland. This is necessary because Scottish legal procedure differs from that followed south of the Border. The annual report is basically statistical in character, *i.e.*. it consists almost entirely of detailed tables apart from a brief 'Introductory note' which is accompanied by summary tables drawing attention to the main features of the year to which the report relates. Examples of these tables are given in the following pages. Apart from the introductory notes the bulk of the report is divided between two main sets of tables, the first of which are 'comparative tables' setting out the data for the period since 1930, only the last two years being given separately, but annual averages for the successive quinquennia facilitate comparison in the changes from period to period in the extent and incidence of crime and convictions. These latter tables give two main sets of figures: first a

[1] See 1 per cent Sample Tables, Great Britain 1951. Part II, Tables X, Nos. 9 and 10.

record of the crimes or 'indictable offences' known to the police, and second a record of the decisions of the various courts, in particular distinguishing between Magistrates' courts on the one hand and the Quarter Sessions and Assizes on the other. The bulk of the tabular material is contained in the third part of the report, consisting of tables setting out in great detail both the police returns of crimes known to them and the records of the courts as well as giving the sentences passed on those convicted. The tables in this last part all relate to the single year covered by the report, *e.g.*, 1956.

A better grasp of the bases of English criminal statistics can be obtained by considering them as prepared on two main lines. The first is an analysis of the persons who have committed crimes and their consequent fate, and the second is an analysis of the crimes committed. The first set of data is derived from the court proceedings and therefore deal only with those cases which the courts adjudge in the year. The second set of data is based upon the police returns. Hence the sub-title of the annual report is 'Statistics relating to Crime and Criminal Proceedings for the year 19—'. Note that as was explained on p.12 no crime exists officially until it has been brought to the notice of the police. And, as was there explained, not all crimes do come to their attention. Even between the different police forces in local authority areas, differences in practice and procedure almost certainly exist. It should also be noted in passing that there are differences in the classification of these two sets of statistics. In

TABLE 38

MALES FOUND GUILTY OF INDICTABLE OFFENCES IN 1956
CLASSIFIED BY AGE AND OFFENCE (PERCENTAGE DISTRIBUTION)

Age group	Larceny	Breaking and entering	Receiving	Frauds and false pretences	Sexual Offences	Violence against the person	Robbery	Other Offences	Total
Age 8 and under 14	64 6	26 0	3 0	0 2	1 5	0 7	0·3	3 7	100
" 14 " " 17	59 8	25 3	4 0	0 4	5 0	3 0	0 4	2 1	100
" 17 " " 21	56 2	22 2	2·6	1·0	4 7	9 1	1 0	3 2	100
" 21 " " 30	57·2	17 3	4 1	2 8	4 9	9 3	0 8	3 6	100
" 30 and over	62 1	6 9	6 0	6·4	8 3	5 8	0 2	4 3	100
All ages	60·4	17 8	4 2	2 8	5·2	5 5	0·5	3 6	100

Source: Criminal Statistics 1956 H.M.S.O.

the first, the crimes are classified on the basis of the conviction, while in the second the police base their classification on the

facts known to them, *i.e.*, the charge. In the event the tow can differ, for example, a case of dangerous driving reduced in the court to one of careless driving, or murder charge reduced to manslaughter.

The statistics cannot be readily understood until the main categories of offence are defined. There are three main classes of what the layman would term 'crimes'. These are 'indictable' and 'non-indictable' together with a numerically insignificant section of 'offences' against the defence regulations. The distinction between indictable and non-indictable offences has become blurred since the classification was introduced some sixty years ago. The basic difference is that the former type of offence can (with few exceptions) only be dealt with in the Court upon indictment, *i.e.*, where the prisoner is accused before a jury. The non-indictable offences (again with exceptions, hence there is some slight overlap with the first type of offence) are those dealt with summarily in the Magistrates' Courts. Of this group motoring offences form about 60 per cent. Within these main groups a standard form of classification is employed. For *indictable* offences there are some 70 different headings set out under six main headings such as Offences against the Person, Offences against Property with Violence, and so forth. This particular classification is justified on the grounds of tradition rather than any especial usefulness. The indictable cases are classified under some 120 types.

TABLE 39

SENTENCES ON PERSONS AGED 21 AND OVER GUILTY OF INDICTABLE OFFENCES BY HIGHER COURTS†

	1938		1954		1955		1956	
	Number	Percentage	Number	Percentage	Number	Percentage	Number	Percentage
Conditional discharge ..	982	15·4	902	7·1	858	7·0	959	7·4
Probation ..	827	13·0	1,530	12·0	1,645	13·4	1,889	14·6
Fine ..	—	—	1,400	11·0	1,453	11·9	1,723	13·3
Imprisonment ..	4,222	66·3	7,664	59·9	7,226	59·1	7,339	56·7
Corrective training	—	—	439	3·4	328	2·7	304	2·3
Preventive detention	—	—	217	1·7	156	1·3	139	1·1
Otherwise dealt with	336¹	5·3	633	4·9	568	4·6	592	4·6
Total* ..	6,367	100·0	12,785	100·0	12,234	100·0	12,945	100·0

Source: Criminal Statistics 1956 H.M.S.O.

¹ This table does not include persons sentenced by quarter sessions after having been found guilty by a magistrates' court.

* Includes fines in 1938 only.

While the work of the Courts is undeniably well and accurately recorded, for purposes of further analysis the existing criminal statistics are considered by criminologists as inadequate. In measuring the incidence of crime, one practice is to relate the number of convictions to the number of persons in each age group in the country as a rate for 100,000 of the population. The age groups, separate for each sex, are annual for ages 8 to 20, and then in 5 year groupings between 21 and 30, and 10 year groupings up to the class 60 and over. Breakdowns of aggregates by sex and four age groups: 8 and under 14, 14 and under 17, 17 and under 21, 21 and over for certain major groupings of offences, and those persons involved in court proceedings, are also given.

For each main category of offence the statistics show the number of offences known to the police and the number of offences cleared up. The overall percentage is usually between 45 and 50 per cent, which at first sight does not suggest a very high rate of detection. But in view of the fact that well over 60 per cent of the 'offences known' are relatively minor cases of larceny, the proportion 'cleared up' of all cases is clearly dependent on the size of that figure. For cases of 'violence against the person' about 90 per cent are cleared up, and in some years well over 90 per cent of fraud cases are settled. A particular weakness of criminal statistics as a measure of the incidence and extent of crime is what is known as the 'dark figure', *i.e.*, the difference between the number of crimes actually committed and the number actually recorded, *i.e.*, brought to the notice of the police. From a statistical point of view this is an important matter, since the proportion of omissions is by no means constant from year to year. Much depends on the willingness of the public to report crimes and their attitude to certain types of crimes, and also to the police, are significant in this respect.

The comment has already been made that the adequacy of criminal statistics in this country is limited from the criminologist's point of view. Their interpretation requires great care. One authority has suggested that in examining these data attention should be concentrated on particular classes of offences at a time. More attention should be paid to the relative fluctuations in the figures from year to year rather than the absolute totals or changes. Only when a persistent trend – up or

down – in the annual totals is observable, is it safe to draw any definite conclusions. For the reader interested in this particular field, a list of references are given below, but for a detailed analysis of the pitfalls for the unwary, the early chapters of Dr. Mannheim's book are especially useful.

REFERENCES

Annual reports on Criminal Statistics for England and Wales, H.M.S.O.
Social Aspects of Crime in England between the Wars. H. Mannheim.
Statistics in Criminology. M. Grunhut. *J.R.S.S.* 1951.
Criminal Statistics. T. S. Lodge. *J.R.S.S.* 1953.

Index of Retail Prices

After long controversy the Cost of Living Index prepared by the Ministry of Labour in 1914 on the basis of the 1904 enquiry into working-class living standards was discontinued in mid-1947 as completely outmoded. Actually, the old index was relatively satisfactory until the outbreak of war in 1939. Then, owing to the concentration of subsidies on its constituent items and its limited coverage, its deficiencies were emphasised, particularly from 1941 onwards, when prices of 'free' goods began to rise more sharply. It was replaced by an *interim* index in June 1947, pending a new enquiry into living standards to be undertaken 'when the expenditure of working-class housewives could be recorded in a market considerably more free' than it was in 1947.

The purpose of the index was to show 'future monthly changes in the level of retail prices weighted according to the pre-war pattern of consumption disclosed by the family budget enquiry of 1937-8'. This enquiry into family expenditures was conducted over a period of 12 months during 1937-8 by the Ministry of Labour. A sample of some 10,700 household budgets kept over four separate weekly periods at quarterly intervals in 1937-8 was finally regarded as suitable for tabulation. They covered three groups of working-class families, *i.e.*, industrial workers, agricultural workers and rural households in receipt of wages up to £250 p.a. in 1937-8. The sample contained over 9,000 budgets of industrial and commercial workers and some 1,500 agricultural workers. The weights for the index were derived from an amalgamation of the two sets of budgets, *i.e.* industrial and agricultural workers, in the proportions of 16 to 1, which represented the approximate proportions of the industrial and agricultural populations in the country.

These budgets provided an analysis of the main forms of working-class expenditure, but ignored the variable weekly outlays on pools, betting, payments for medical treatment, and insurance premiums. They also provided the basis of weighting

TABLE 40

Group	Weights Interim Index of Retail Prices 1952	Weights Interim Index of Retail Prices 1947	Cost of Living Index 1914
i. Food	399	348	60
ii. Rent and Rates	72	88	16
iii. Clothing	98	97	12
iv. Fuel and Light	66	65	8
v. Household Durable Goods	62	71	4
vi. Miscellaneous Goods	44	35	
vii. Services	91	79	
viii. Alcoholic Drink	78	101 } 217	
ix. Tobacco	90	116 }	
	1,000	1,000	100

the individual items in the various groups of expenditure, which are contrasted in Table 40 with the proportionate weights employed in the discontinued 1914 Cost of Living Index and with the revised 1952 weighting. These budgets provided an interesting illustration of the difficulty of obtaining accurate data. The Ministry of Labour in analysing the budgets found that there were indications that the household expenditure on tobacco and cigarettes 'was not in all cases fully stated' and that personal expenditure on beer, spirits, etc. 'was not fully reflected in the data provided by the survey'.[1] It was consequently necessary to employ other data to determine the appropriate weighting to be given these particular items.

The more detailed classification of the 1947 index indicates its greater coverage. A sub-group of the 1914 'Other Items' (not shown in the table) was given four times as heavy weighting in the 1947 index as in the 1914 index, while nearly one-quarter of the 1947 weighting (*i.e.* the goods to which it related) was not represented in the 1914 index. They cover articles and services which did not enter into the 1904 survey such as petrol, postage, motor and wireless licences, laundry, hairdressing and alcoholic drink. The totals 100 : 1,000 do not imply that one set of weights

[1] Interim Index of Retail Prices. Method of Construction and Calculation (Revised Edition) H.M.S.O. 1952.

are ten times as large as the other; the 'weighting' is relative as between the constituent items. If the reader wishes, he may add '0' to all the 1914 weights. The results, if applied to actual data, will be exactly the same as without them.

The base date selected for the original Interim Index was the 17th June 1947. As its name implied, the index was to be used only until such time as a new index could be constructed on the basis of post-war data relating to consumer expenditure. In June 1951 the Cost of Living Advisory Committee recommended the holding of a new enquiry on household expenditure to provide

TABLE 41

INTERIM INDEX OF RETAIL PRICES*

Group	17th June 1947 = 100		15th January 1952 = 100		
	1950 17 Jan.	1952 15 Jan.	1953 13 Jan.	1954 12 Jan.	1955 18 Jan.
(i) Food	120·3	149·7	109·2	110·2	119·2
(ii) Rent and Rates .	100·4	104·2	103·7	109·9	113·5
(iii) Clothing . ..	117·1	147·1	94·9	96·1	95·6
(iv) Fuel and Light ..	115·1	140·1	104·5	110·7	114·9
(v) House Durable Goods ..	108·1	136·6	97·6	95·6	95·3
(vi) Miscellaneous Goods ..	113·6	137·3	102·7	100·0	99·4
(vii) Services	106·1	123·9	107·1	109·5	112·6
(viii) Drink and Tobacco ..	107·5	108·5	101·0†	101·5†	102·5†
All Items	113	132	104·4	105·8	110·2
Old and New Indices linked to base 1947 = 100	113	132	138	140	146

* Ministry of Labour Gazette.

† Drink only. The tobacco relative was 100·3 throughout this period.

the basis for a new and more permanent index. The same committee in a later report made a number of recommendations for modifying the basis of the Interim Index until such time as the new index could be started.[1] The changes recommended by the committee concerned mainly the weighting employed, and the new index with the revised weighting was based on January 1952. It will be noted from Table 41 that it has been possible to link the old and new 'all-items' indices, but it is not possible to link the section indices because of the revised weighting. The introduction of the revised index with the changed base was

[1] Report on the Working of the Interim Index of Retail Prices, Cmd. 8481. See also earlier reports Cmds. 7077 and 8328. H.M.S.O.

justified as a result of the changes in the pattern of consumption in the post-war period from that of 1937-8. The Cost of Living Advisory Committee concluded its enquiry on the operation of the Interim Index by arguing that the 'all items' index was a fair representation of the change in prices that had taken place since June 1947. It was admitted that the index was unsatisfactory in respect of rents of new houses built since 1947 which were generally higher than the old; but the extent to which the index of 'all -items' was understated thereby was offset by the tendency of the pre-war pattern of consumption to overstate the rise in food prices.

In January 1953 a new enquiry was begun to obtain an up-to-date basis for a revised index. It will be remembered that the last comprehensive enquiry into household expenditure took place in 1937-8. The enquiry 1953-4 was based upon a multi-stage sample of 20,000 households throughout London and 350 other areas including a proportion of rural districts. The sample was designed to yield a representative cross section of all households including the large family, the childless couple, and pensioner living on his own. A sample of 8,000 households would have been quite adequate for the needs of the survey, but the 20,000 names were drawn to ensure that on the basis of a probable response rate of 40 per cent, the required 8,000 would be obtained. The estimate of non-response was based upon experience with the National Food Survey over a number of years. All persons within the 20,000 selected households were asked to keep a record of their expenditure over a period of three consecutive weeks. Different households had to keep their records for different periods, so that the whole of the year was covered by the enquiry. In other words, there was an even flow of interviewing and recording throughout the year in each main region of the country. All members of the household aged 16 or over who had left school were required to submit a confidential return of his or her personal expenditure. Apart from the data so obtained the interviewer obtained facts concerning the household from the householder or housewife; *e.g.* whether the dwelling was rented or owned, and the costs incurred upon the house. Each adult member of the household was asked to give information about his (or her) expenditure (generally covering a year) on annual licences, motor tax and insurance, insurance

other than motor, education and training fees, season or contract tickets. All this information was required in addition to the completion of the three weekly statements of expenditure. Each member of the household providing these details was paid £1. From the foregoing it is clear why so low a response was to be feared, but in the event the results far exceeded expectations. Almost 65 per cent. of the households co-operated to yield a total of 12,911 budgets.

The Cost of Living Advisory Committee reported the results of the enquiry on household expenditures in March 1956. Their 'Report on Proposals for a new index of Retail Prices' (Cmd. 9710) recommended a new index of retail prices with a base date January 1956 = 100. The data for this new index and its weighting was derived from an analysis of 11,638 budgets which, in the opinion of the Committee, 'will provide a reliable and representative basis for calculating the weighting pattern for a new index'. The actual number of budgets received from the households sampled – nearly 20,000 – was 12,911 and the decision to use the smaller number of 11,638 was based upon the elimination of two small groups. These were the households in which the head of the household received a recorded gross income of £20 per week or more and those in which at least three-quarters of the total income was derived from the National Insurance retirement pension and national assistance. The reasons adduced for excluding the first group of 460 households was that the pattern of expenditure of such households differed from that of the majority of households in receipt of smaller incomes. Furthermore, there was much greater variability between such households which makes any average rather unsatisfactory. As for the second group containing 813 households with 1,216 persons (nearly 60 per cent. of such households consisted of one person living alone) it was decided that this fairly homogeneous group should be excluded since its pattern of expenditure also differed appreciably from that of the average household covered in the survey. According to the Committee, 'the intermediate group of 11,638 households (from which budgets were collected) will provide a satisfactory foundation' for the new index. This sample of households would 'reflect the expenditure pattern of nearly nine-tenths of all households in the U.K.'

It seems therefore reasonable to conclude that the new index is 'representative' of that section of the community whose cost of living it is designed to measure; or more precisely, the changes in its cost of living. The Committee introduce a few caveats on this point. It draws attention to the fact that as in the 1937/8 enquiry, the expenditure recorded by the households in respect of alcoholic drink and tobacco was substantially under-recorded. The Committee was able to estimate the appropriate weights to be assigned to these items from other information, *e.g.* Customs and Excise data. Similarly, there was some under-recording of expenditure on meals consumed outside the home, on sweets and chocolate as well as on ice cream and soft drinks. Apart from providing the details of their weekly outlays, members of the households constituting the sample were asked to state their incomes, and 96 per cent of the informants did so. The purpose of this information was not so much to provide information relating to the construction of the new index as to provide a breakdown of the households by income levels. Unfortunately, the marked willingness to cooperate in this matter was not accompanied by an equally strict regard for accuracy. The Committee's report states that 'in very many cases they (the informants) must have considerably understated their incomes'. This, explains the Committee, is partly

TABLE 42

Group	Old Weights (1950 Consumption)		New Weights (1953-54 Consumption) at January 1956 prices
	At January 1952 prices	At January 1956 prices	
Food	399	432	350
Alcoholic Drink	78	69	71
Tobacco	90	80	80
Housing	72	73	87
Fuel and Light	66	73	55
Durable Household Goods	62	55	66
Clothing and Footwear	98	84	106
Transport and Vehicles } Services	91	94	126 { 68 / 58
Miscellaneous Goods	44	40	59
ALL ITEMS	1,000	1,000	1,000

Source: Report on Proposals for a New Index of Retail Prices. Cmd. 9710.

explained by the tendency to regard the 'take-home' pay as constituting the income, *i.e.* income after tax, insurance and similar deductions. There was, however, frequent omission or understatement of overtime pay, bonuses and earnings from subsidiary employments etc. The Committee charitably states that this was 'in many cases unintentional'. Despite these imperfections, however, the Committee is confident that it has obtained a 'reliable and representative collection of budgets'.

The classification of the expenditures comprising the index and the corresponding weights are given in Table 42. The groups, ten in number, are not very different from those used in the 1947 and revised 1952 indices illustrated in Table 41. The order has been slightly changed and there is one new group, 'Transport and Vehicles'. The new weighting is given in the last column of Table 42 and should be compared with the original weighting of the 1952 revised interim index in the first column. The second column reflects the effect on the weighting if the 1952 index (which was based upon the estimated consumption pattern in 1950) were re-weighted to take account of the effect of price changes between 1952 and 1956 on the 1950 consumption pattern. The weights to be used in the new index differ from the latter because they are based not upon the 1950 consumption pattern, but upon that in 1953/54 as revealed by the household budgets. In any case, statistically speaking the two indices, *i.e.* 1952 and 1956 are not comparable because the coverage, *i.e.* range of households, of the later index is much broader than the 1952 index which was based upon the data obtained in the 1937-38 enquiry. In short, *the index based on 10th January 1956 = 100 is a new index which cannot be directly compared with its predecessor*. Nevertheless, the Committee recommended that for a while the 'All items' index of the 1952 interim index should be linked with the corresponding index for 1956. It will be remembered that a similar procedure was followed with the introduction of the 1952 interim index, so that it will still be possible to measure roughly the broad change in retail prices since 1947

The apparent similarity of the weighting of the 1952 and 1956 indices tends to obscure the fact that the new index is based upon a very different pattern of expenditure. Essentially, the 1952 index was still largely based upon a pre-war pattern. In

the Appendix to the report of the Advisory committee the constituent items of each group, together with the appropriate weight, are given. There are many additions, in particular a much wider range of foodstuffs; the alcoholic drink group now includes spirits other than whisky. The durable household group includes the television set and washing machine as well as a dining room suite which had not appeared before. Clothing now includes a made-to-measure suit and pyjamas, while the nylon slip and fully fashioned nylon seem to have replaced the flannel petticoat and lisle hose of 1914.

The reader may recall that in selecting the constituent items of any index, care has to be taken to ensure that they are representative of the 'population' of similar items. In the Retail Price index there are 91 sections and within each section there is a number of selected items of expenditure. According to the report of the Advisory committee, 'the principle followed in selecting these items has been to choose items whose price changes can be regarded as representative of the average change in the prices of the whole range of goods covered by the weight assigned to the section in question'. The report goes on to explain that although only a few items are included in each section, the weight allotted to the section is determined not by the total expenditure on the selected items but by that on all items of that type. The average price changes shown by the selected items are 'regarded as representing the average price changes for all kinds of goods covered by the relevant section'.

The calculation of the index requires the computation of the price relatives at monthly intervals for each item of expenditure in each town from which the data are obtained. Information concerning food prices is generally collected by personal visits to retailers by local Ministry of Labour Officers; usually from five retailers considered by the manager of the local office to be typical of the shops where working-class households normally do their shopping. Prices are obtained in respect of 200 local office areas grouped according to the population of the towns as follows:

A.	Greater London	25 areas
B.	Towns with a population of 200,000 and over	25 areas
C.	Towns with a population of 50,000 – 200,000	50 areas

D. Towns with a population of 5,000 – 50,000 .. 50 areas
E. Towns with a population of under 5,000 .. 50 areas[1]

Some prices are communicated directly by the manufacturers to the Ministry, *e.g.*, clothing prices. For those commodities where quality changes are significant, *e.g.* as with clothing, the relatives of current prices are calculated not against the base price, but on the chain-base system.

The actual calculation of the index is briefly as follows. The A.M. of the price relatives for each town or area is averaged with those from other towns and areas with due regard to the size of the area's population, in order to obtain the overall national index which relates to prices on the Tuesday nearest the 15th of each month. The official account sets the process of calculation out in detail as follows:

(A) The price relative for each item in each town is calculated and the resulting figures combined as an unweighted average for all the towns in *each* population group.

(B) The separate indices for the various population groups are averaged to give indices for each *item* for the country as a whole.

(C) The national indices of the items, *e.g.* bread, are next combined to arrive at indices for each group, *e.g.* food. In the construction of the group index, the percentage increase in each item in the group is weighted by reference to its proportionate share in the aggregate outlay on all items in that group. Thus the group index is the weighted arithmetic mean of the percentage changes in its constituent prices.

(D) The indices for the various sections are then combined, being weighted as shown in Table 58 on page 340 which gives the index for each of the main expenditure groups, I to X, and the final all-items figure. It will be noted from Table 58 that all indices for sections are given to one place of decimals. This has the advantage that quite minor changes in prices will be reflected in the index.

The student should note that the Committee recommends a modification to the method of collecting prices of clothing, footwear and household textiles. At present, the price changes are

[1] The towns actually surveyed are given in an appendix to the official account of the Index. The grouping by population and the selection of the localities within each population group are made in such a way as to give adequate representation to different types of localities throughout the country

calculated for four types of area; London, other very large cities, intermediate sized towns and smaller towns. The changes in each area's price relative are averaged, each being equally weighted. The Committee felt that with the growth in importance of the multiple and department store, the collection should not be by geographical area, but by type of shop. In this case the four groups would be multiple stores, department stores, co-operatives and other shops. The weights for these four groups' price relatives should be based upon their relative importance in the national sales of clothing. This would entail the collection of data relating to the value of such sales for the last two groups of shops.

The item of Rent and Rates was also considered at length by the Committee. Owing to the greater flexibility in rents following their de-control and the emergence of the Local Authorities as landlords on a large scale, new information is needed. The Committee recommends that a sample of the households co-operating in the household expenditure enquiry should be visited at regular intervals to ascertain whether any changes in the rental or rates has taken place. The change recorded in this item for the sub-sample would then be assumed to be indicative of the average change in the rents of rented dwellings. The Report states that about 30 per cent. of the budgets came from households owning the dwellings in which they live; roughly one half of their number is in the process of buying the dwelling. This expenditure is regarded as a form of saving and as such does not enter into the new index, but the Committee recommends that an estimated rental should be imputed to such dwellings and it should be merged with the data relating to rented dwellings for purposes of determining the weight to be allocated to the Housing group within the index. In this group, account is taken of the expenditure on house repairs and decoration.

A report published in 1957 contains a detailed account of the 1953/4 enquiry upon which the current index of retail prices is based. Not merely is the planning and conduct of the survey explained, but the report discusses in detail the various problems arising from the interpretation of the expenditure schedules provided, *i.e.*, an assessment of the accuracy of the data supplied. The bulk of the report is then devoted to a detailed tabular analysis of households classified by income, by their com-

position, by region, etc., and the various expenditure patterns of different income groups are contrasted. Table 7 which is reproduced on page 45 is a small sample of the type of information given. The report is a mine of information for the study of living standards in this country at the present time.*

REFERENCES

For a summary account of the enquiry and exposition of the current index see the 'Report on Proposals for a New Index of Retail Prices' Cmd. 9710. A detailed account of the actual construction of the index is given in 'Index of Retail Prices. Method of Construction and Calculation'. This is part of the Industrial Relations Handbook (H.M.S.O.) but is also published separately as a pamphlet.

*The above mentioned report on the 1953-4 enquiry is known as the 'Report of an Enquiry into Household Expenditure in 1953-54' H.M.S.O. 1957.

Wages and Earnings

Statistics of income in the United Kingdom are available from two main sources. The first is the Annual Report of H.M.Commissioners of Inland Revenue which provides information such as that contained in Table 6 (page 44). In addition to such classification of incomes according to size, there is an analysis of incomes by region and some interesting breakdowns of income groups by reference to the personal allowances claimed, *i.e.*, single, married, one two and more children, dependent relatives, etc. Such data are, of course, of interest only for the reader requiring an overall picture of the distribution of income. The Blue Book on National Income and Expenditure supplements the latest annual analysis of taxpayers by income size, based on the Inland Revenue data, with additional tables for earlier years.

The most important and comprehensive source of statistics on wages and earnings is the *Ministry of Labour Gazette*. These data relate to manual workers in a wide range of industries and various classes of worker, *e.g.* male and female, adults and juveniles. They have been published for many years. Some information was collected officially as early as the mid-nineteenth century, while statistics of wage rates and hours of labour were published separately in book form from 1893 onward. Originally the publication of this book was at irregular intervals but since 1946 this volume, known as *Time Rates of Wages and Hours of Labour*, has appeared annually. The data contained in that volume relate to the rates and hours applicable to manual workers in each industry at the beginning of April each year. To keep the information up to date any changes, which arise from

voluntary collective agreements between employers and the unions within the separate industries, are published in the monthly *Gazette*. The Ministry of Labour has no compulsory powers in regard to the collection of these data, but voluntary co-operation from the unions and employers' representatives is, according to the official report, 'very freely accorded'.

The data so provided and published in this annual handbook serve as the basis of two official index numbers. The first is an index of weekly *wage rates* which is designed to measure the average movement from month to month in the level of full-time wages in the principal industries and services in the United Kingdom for the main groups of workers, *i.e.* adult males and females and juveniles, as well as an 'all workers' index. The index was actually started at the beginning of this century but has since been revised six times; the present base is 31st January 1956.[1] Before that date, there is a single continuous series running from June 1947 to January 1956, with June 1947 as the base date. It should be noted that the new base corresponds with that for the index of retail prices discussed in the previous section. Owing to the

TABLE 43

INDEX OF WEEKLY WAGE RATES*

All industries and services†

(31st January 1956 = 100)

Period	Men	Women	Juveniles	All Workers
1956 Monthly Averages	104·8	104·2	105·5	104·7
1957 Monthly Averages	110·0	109·7	113·3	110·0
1958 Monthly Averages	113·8	1140·0	115·8	113·9

* *Source: Ministry of Labour Gazette, February 1959*

† A similar group index is compiled for the 'Manufacturing Industries only'.

changes in the data available, as well as alterations in weighting, it is not possible to compare successive series of index numbers of weekly rates of wages calculated on the different base years. It is possible, however, to make an estimate of the overall change in 'all workers' wage rates between September 1939 and 31st January 1956. Note that this index (or indices since one is prepared for each of the three categories of employee as well as 'all workers') is based upon weekly *rates*, not earnings. Thus it will

[1] The method of constructing this index was discussed in the *Gazette* for February, 1957.

show no change if workers' earnings alter when extensive overtime is being worked, or if a large proportion of the labour force is on short time. The current index is based upon the rates agreed as at 31st January 1956, and the weighting of the index for each class of worker in each industry is determined by the relative size of the total wage bill. The index for each class of worker for the past three years is given in Table 43.[1]

Another index on the same base date is calculated to measure changes in the normal working week. This index is known as the index of *normal weekly hours*. By dividing the index of weekly wage rates by this index of normal weekly hours, the Ministry of Labour produces what it calls an index of *hourly* rates of wages. It will be appreciated that this index reflects any improvement in the manual worker's terms of work, since any reduction in the length of the base week (without any change in the wage for the basic week) represents an improvement in the rate of earning.

Quite apart from the statistics relating to wage rates, the Ministry of Labour publishes extensive data relating to *earnings* and *actual* hours worked. These are, of course, quite distinct from the figures discussed above relating to wage *rates* and *normal* hours worked. This distinction must be kept in mind when consulting the published data, since it is easy to confuse the different series. As with the data on hours and rates, the statistics of earnings and hours worked go back to the last century, although there have likewise been frequent changes in the basis of their compilation. The current series date from 1938 and their purpose is to provide information relating to the general trend of actual gross earnings and weekly hours worked. The statistics published include tables such as Table 44 below showing the average weekly earnings for each class of worker in various industries. Similar tables, *i.e.*, based on the same classification, are published in the same issue of the *Gazette*, showing average hours worked and average hourly earnings for each of the four age and sex groups for a wide range of industries. In addition to the figures for individual industries, comparable figures are calculated for industrial groups as well as for 'all industries' combined. The actual basis of the published figures is described in the half-yearly articles in the *Gazette* and these should be studied before figures are extracted.

[1] The principle of weighting in index number construction is explained in Chapter XVI.

The half-yearly enquiry (see February and September *Gazette*) into earnings and hours of work of manual workers covers 7 million workers in over 68,000 establishments, excluding agriculture, mining, railways, dock labour and the service trades: equal to rather more than two-thirds of the total number of wage-earners employed in the industries covered by the enquiry. In the enquiry of April 1958 employers were asked to make returns showing the number of wage-earners at work in the last pay-week in April, of the aggregate earnings of those wage-earners in that week, and of the total number of man-hours worked in the week; these various data being classified under the following headings: men 21 years and over, women 18 years and over, youths and boys under 21 years and girls under 18 years. Table 44 illustrates the data compiled from such an

TABLE 44

AVERAGE WEEKLY EARNINGS IN THE LAST PAY-WEEK IN APRIL 1958

(Manufacturing industries only)

Industry Group	Men (21 years and over)		Youths and Boys (under 21 years)		Women (18 years and over)		Girls (Under 18 years)		All Workers	
	s.	d.	s.	d.	s.	d.	s.	d.	s.	d.
Treatment of non-metalliferous mining products other than coal .. .	251	1	120	8	122	10	82	9	217	6
Chemical and allied trades ..	258	0	115	1	128	3	84	0	223	7
Metal manufacture . .	280	2	124	8	138	0	89	0	261	10
Engineering, shipbuilding and electrical goods .. .	264	11	104	2	141	8	88	2	229	3
Vehicles	281	3	100	2	150	2	92	10	251	7
Metal goods not elsewhere specified	264	2	106	11	129	0	83	3	212	3
Precision instruments, jewellery etc.	253	5	102	0	135	7	83	4	199	1
Textiles	228	11	108	2	129	8	93	11	166	10
Leather, leather goods and fur..	225	1	105	8	124	1	79	4	174	1
Clothing	223	7	103	10	127	2	82	7	143	3
Food, drink and tobacco ..	236	7	107	7	126	9	85	2	186	8
Manufactures of wood and cork	233	5	104	10	135	6	79	0	202	8
Paper and printing . ..	289	11	111	5	136	9	81	7	228	0
Other manufacturing industries	258	8	114	4	128	8	85	6	203	8
All manufacturing industries ..	261	4	106	10	131	8	85	11	211	11

Source: Ministry of Labour Gazette, September 1958.

enquiry. It is important to note that these data exclude office workers, shop assistants, outworkers as well as salaried staff. The wages returned are the total earnings, inclusive of bonuses, before tax and insurance deductions. Although this enquiry is dependent upon co-operation of employers, of 69,400 establishments to which forms were sent, 68,200 of the returns made were suitable for tabulation. The data obtained from this enquiry are too detailed to be reproduced here and the student is advised to consult the *Gazette*, *e.g.* September 1958, to ascertain the nature of the published information. Whenever these tables are published the reader is warned against careless interpretation of the figures given. 'In view of the wide variations, as between different industries, in the proportions of skilled and unskilled workers, in the opportunities for extra earnings from overtime, night-work and payments-by-results schemes. . . . the differences in average earnings shown should not be taken as evidence of, and a measure of, disparities in the ordinary rate of wages prevailing in different industries for comparable classes of workpeople employed under similar conditions.' In other words, for each of the four categories of employees, the earnings shown are *averages*, and as the student has already learnt, the mean is not always a reliable guide to the data. In this case, the differences in earnings between various industries are not completely explained by the differing *rates* of pay in these industries.

REFERENCES

Labour Statistics. Guides to Official Sources, No. 1, H.M.S.O., 1958, Revised Edition.

Conclusion

The student reader must realise that the foregoing summary can only provide an indication of the main sources of some branches of social statistics. It cannot be sufficiently emphasised that a real understanding of these published data can be acquired only by studying the publications, and actually extracting data relating to a particular problem. For example: how many families in your home town were living in 1951 more than two persons per room? To what extent is it true that crimes of violence have increased since the war? By seeking in the relevant publications answers to such questions as these, the student will learn more than from reading this chapter a dozen times over!

CHAPTER XV

VITAL STATISTICS

Introduction

In recent years the term 'vital statistics' has been defined as the use of three figures to describe a single figure. Popular usage apart, vital statistics are derived from the enumeration of the human population and the registration of births, marriages and deaths. The statistician specialising in this branch of statistics is usually described as a demographer, the subject of demography being the analysis of population data. The collection of information about the human population dates back to Biblical times. The Romans used such information for raising both armies and tax revenues. Nevertheless, collection of vital statistics in the United Kingdom on the systematic basis of present times is little more than a century old, although the first official English population census was held in 1801. With the creation of the General Register Office in 1837 there was instituted a continuous system of registration of births, marriages and deaths. The G.R.O. was also made responsible for the decennial census, starting with that of 1841.

While the primary purpose of statistics relating to the human population in a country was initially military and fiscal, the main impetus to the improvement of the system in the nineteenth century came from the need to improve public health. As far back as the reign of Henry VIII, weekly Bills of Mortality had been compiled to provide information concerning the spread of the plague and the passage of the Act setting up the General Register Office was almost certainly expedited by the experience of the cholera epidemic of 1831. The growth of the new industrial areas in the mid-nineteenth century and the attendant high mortality therein led to the demand for accurate registration of both the populace of such towns and the number of deaths as well as their cause, although it was not until the 1890's that compulsory notification of certain diseases was introduced in the United Kingdom.

Throughout the first century of the G.R.O.'s existence its primary concern was the collection and analysis of data drawing attention to conditions of public health in the industrial towns. Even at the present time the annual reports of the Registrar General devote considerable space to such matters as epidemics and differential mortality, but since the end of the first World War the attention of demographers has been concentrated on population trends. The tremendous rate of growth in the population of most countries during the middle and latter half of the last century ended before 1900 and in recent decades concern has been expressed regarding the contraction in the size of the average family. In the mid-1930s various forecasts were made regarding the future size of the United Kingdom population, many of them depressing and some quite alarming in their implications. It was this concern with the trend in family size which led to the setting up of the Royal Commission on Population in 1944. Its report, which was published in 1949, as well as the post-war increase in the number of births, has led to the rejection of the more pessimistic pre-war forecasts regarding the trend of the British population. But interest in the demographic trends elsewhere in the world, in particular the dangers of overpopulation in the Middle and Far East have prompted demographers to analyse and consider their implications for the long term future of these areas.

The demographer is thus concerned with two aspects of vital statistics. There is first the collection and tabulation of current data relating to births, marriages, sickness and death and their analysis. From these data it is possible to ascertain the changing causes of death over long periods and the need to develop new medical techniques to cope with special problems. For example, at the present time it is apparent that the main hope for reducing the death rate among infants under one year of age lies in reducing the risk of premature births and deaths in the first month of life rather than in the following eleven months. It is primarily the study of such data and the consequent action taken on the basis of such findings in the form of public health services and better living conditions that has led to the male expectation of life at birth rising from 40 to 66 years between 1851 and 1951. The second aspect of vital statistics deals with the analysis of population trends, in particular the attempts to forecast its future

size and pattern of change. The purpose of such analyses is to ensure that the appropriate administrative action in respect of housing, schools and other social policies will be taken. The pre-war policies of a number of West European countries to encourage more births were introduced following upon such demographic analysis.

The information comprised in the term 'vital statistics' is collected in England and Wales by the General Register Office through the medium of its 1200 local offices. Scotland has its own G.R.O. The data are collected in two ways; by means of a continuous system of registration and by means of a periodic census which in this country is held every ten years. It would be possible to derive some vital statistics by either method alone, but by using both methods a periodic check is provided on the data collected and any analysis of such data is in consequence more reliable.

The population census has been held regularly in the first year of each decade between 1801 and 1951, the only omission being in 1941. Some limited data for this period are provided by the National Registration figures of September 1939. It is usual to hold the census at a time of the year when the movement of the population is expected to be fairly low. If the census were to be held on August Bank Holiday, the seaside towns would show vastly inflated populations and the urban areas correspondingly reduced numbers. To minimise this problem recent census enumerations have been made on a Sunday in early April.

The census entails a simultaneous enumeration of all the households in the whole country, *i.e.*, the United Kingdom, on one day in the year. Each head of a household, and there are some 11 million in England and Wales, is required to fill up the census form giving details of all individuals resident under his roof on the cenus night. This method of enumeration is known as a *de facto* census, *i.e.*, a count of the population present at the place and time the census is taken. In the United States the census is on a *de jure* basis. In this case the head of the household lists the members of his household whether they are present or not. The advantage of the *de jure* basis is that it gives a complete picture of the population on a basis of their usual residence but it suffers from the fact that both omissions and double-counting of absent-members is possible, a weakness from which the

English method is free.[1] Census returns are also made by Institutions, Hotels and Boarding Houses. The forms are then collected by locally appointed enumerators who scrutinise the forms on collection for omissions and occasionally help the householder complete the form. After further checking the forms are sent to a central office for recording and tabulation. The use of machine tabulation and the introduction in the 1951 census of a sampling procedure enabled the G.R.O. to publish in less than two years the two volume report on the 1 per cent sample in Great Britain which constituted a census in miniature.

It cannot be emphasised strongly enough that the population census is much more than a mere counting of heads. In the 1951 census questions asked related to the age, sex, marital status, residence, education, birthplace and nationality, occupation and place of work of each individual in the household, as well as questions relating to the household arrangements for water supply and cooking. The census provides an instantaneous photograph of the community at a given point of time in respect of its social and demographic condition. The census is also valuable because it permits the government to collect information regarding specific social problems which are important at the time of the census. For example, the 1911 census provided information relating to fertility, that of 1921 yielded data on the subject of dependency as a preliminary to the introduction some years later of a pension scheme for widows and orphans. In 1951 questions on fertility were again asked to supplement the 1946 enquiry on this topic while new questions on housing amenities were also included.

It would be idle to pretend that all the information collected at the decennial census is accurate. There are two main sources of error. The first source derives from the administrative arrangements. It would be fanciful in the extreme to believe that an *ad hoc* corps of over 40,000 willing but part-time enumerators can contact *every* household or avoid occasionally mis-directing those householders who seek their advice, or they may insert inaccurate information on the form where its completion is the responsibility of the enumerator. Errors also arise at the clerical stage when the information on the schedules is transferred on to

[1] Note, however, that the English census is taken on a day or two, but in the States the enumeration is spread over as much as four weeks, but the composition of the household is recorded as at a given hour.

the tabulations via punched cards. In the course of such work, mistakes in classification are also unavoidable, although they can be kept to a minimum with proper supervision. The second and more important source of census errors is the informant himself. These are of two kinds, wilful misstatements of fact and genuine mistakes due to ignorance or misunderstanding of a question. Of the former, the best known is the unwillingness of many people to state their true age. Analyses of the final data reveal that there are marked concentrations in any age distribution of the population about the ages ending in 0 and 5, while the number of women between the ages of 20 and 30 is overstated due to the desire of many to withhold the fact that either she is still in her 'teens or has passed into the 'thirties'. In the case of older people ignorance or poor memory may account for the fact that there is a tendency to over-state the age after retirement. Divorced women are often reluctant to disclose the fact, describing themselves as single or widowed, although with the changing social attitude towards divorce this may no longer be so serious a source of error as it has been in the past. One particular question on the census form which has always given great difficulty is that asking the occupation of the individual. There is a tendency to over-state the status of the post held, while retired and unemployed persons often described themselves as members of the trade in which they used, or if unemployed, customarily worked. Some concern has also been expressed by the census authorities regarding the accuracy of answers given to the questions on education contained in the 1951 census, which it appears, were not always comprehended by the informant.

The Registration System

The decennial population census is supplemented by a system of continuous registration in over 500 local areas of the United Kingdom. The principle of registering certain vital events has long been established; certainly long before the need for periodic censuses in this country was generally accepted. It is noteworthy that refusal to complete the census return has still to be discouraged – penalties being imposed on the head of the household. Before the G.R.O. many such records were kept by the Clergy. The justification for such records is derived largely from legal considerations, for example the need to prove relationship

to the deceased to benefit under a will. We are all familiar with the birth certificate and the need for its production at certain times, *e.g.*, on taking out a life insurance policy. Nowadays, social and economic needs are just as important a justification for the expense and trouble of the census. The Act of 1836 set up the G.R.O. and many local offices at which any birth, marriage or death had to be notified within a specified period. Another Act in 1874 provided for each death registration to be accompanied by a medical certificate stating the cause of death. Before this date, detailed analysis of mortality statistics was not practicable. Another gap in this country's demographic statistics was closed by the passage in 1938 of the Population (Statistics) Act. This was enacted to provide fuller information on fertility. Additional details were required on the registration of a birth apart from the names of the child and its parents. In particular the age of the mother, the duration of her present marriage and number of other children by the present and any previous husband were recorded by the Registrar. These data are not entered in any public register but are confidential to the G.R.O. By the same Act, on registering a male death, his marital status must be disclosed, while in the case of deceased married females, the year and duration of the marriage are needed together with information regarding the number of children by the present husband whose age at date of wife's death is also required by the Registrar.

The accuracy of such statistics is determined by efficiency of the registration system. In this country accurate and prompt registration is enforced largely by law and habit. To bury the deceased, a copy of the death certificate must be produced to the undertaker. In the case of marriage there are equally obvious sanctions, although, of course, registration is effected during the ceremony. Only with respect to births is there any doubt, for a birth may be registered up to 42 days after the event. Before the war, the birth statistics were classified by calendar quarters according to the date of registration, and the average time lag between the birth and its registration was one month. With the introduction of the war-time food rationing scheme, a new birth meant a new ration book and to claim this prior registration was necessary. The time lag largely disappeared and since the beginning of 1941 the Registrar-General has classified births by

the date of their occurrence instead of by date of registration. The birth statistics of the two preceding years 1939 and 1940 have also been published on the revised basis.

Some reference is needed to the large volume of published material produced by the General Register Office. The data collected by the process of registration is published annually in the form of a three-part report known as the Registrar General's Review of England and Wales. The three parts are known as the Medical, Civil and the Text or Commentary volumes. The last of these appears with a greater time lag than the first two. The *Medical* volume is concerned with mortality and provides detailed tabulations of the causes of death at different ages of life for various groups of the population such as males, females and infants. The *Civil* volume is primarily concerned with population counts, births, marriages, and in recent years with divorce, the statistics of which are still limited. In the same volume there is contained most of the official information on fertility. The final Text, or *Commentary* as it is now known, is devoted in the main to a commentary on the events recorded in the tables of the first two volumes and in particular with changes in the trend over longer periods.

There is also a weekly return of births, deaths and notifications of infectious diseases in all administrative areas of England and Wales and the 10 'standard' regions into which the country is divided for government purposes.[1] These are compounded into a quarterly return which also includes marriages and a brief commentary on any significant features revealed in the tables. These returns form the basis of the Annual Reports. Since 1948 the G.R.O. has also prepared mid-year estimates of the population of each administrative area in England and Wales. These are needed for various administrative purposes by the Central and Local Governments.

The publications of the decennial census are too numerous to enumerate here; they will be found listed in any single volume of the Census Report. They appear at intervals of months and years following the census. Within a few months of the census a Preliminary Report is issued which contains both an account of the conduct of the census and the new population figures for each Local Authority area compared with those of the previous

[1] As shown in Table 10, page 49

census, indicating the absolute and relative changes in their populations. As part of the 1951 census a two volume report based on a 1 per cent sample was produced; this was described above as a census report in miniature. This experiment, designed to expedite publication of the main features of the census, was tried out for the first time in 1951 and it is highly probable that it will be repeated in the future. The County volumes appear next and at intervals the various volumes on particular topics such as Occupations, Housing, etc. The last volumes to appear are the General Tables and the General Report; these constitute a survey of the main findings of the census.

The data collected during the census also forms the basis for certain special studies by the Registrar General which are only practicable given the most accurate data available. They are described as the *decennial supplements* and of these the most important are the *life tables* and analyses of *occupational mortality*. In passing, reference should be made to the publications of the Royal Commission on Population, in particular the important study of fertility conducted by Professors Glass and Grebenik. A list of references is given at the end of this chapter.

Having reviewed the main statistical information collected by the G.R.O. we must now consider the various uses to which these data are put and the form in which they are published so that they may be correctly interpreted.

TABLE 45

POPULATION OF ENGLAND AND WALES 1841–1957

Census Year	Population 000's	Absolute Increase 000's	Decennial Percentage increase %
1841	15,914		—
1851	17,928	2,014	12·7
1861	20,066	2,138	11·9
1871	22,712	2,646	13·2
1881	25,974	3,262	14·4
1891	29,003	3,029	11·7
1901	32,528	3,525	12·2
1911	36,070	3,542	10·9
1921	37,887	1,817	5·0
1931	39,952	2,065	5·5
1941*	41,748	1,796	4·5
1951	43,758	2,010	4·8
1957*	44,907	1,149	4·4†

* Mid-year estimates; these were not census years.
† Adjusted on to a decennial basis. *Source: Annual Abstract of Statistics.*

Table 45 provides some basic data relating to the trend of the population of England and Wales since 1841. Although the first census was taken in 1801, it was not until 1841 that the newly created G.R.O. was made responsible for the census. The earlier census totals were undoubtedly too low due to omissions, and comparable figures exist effectively only since 1841.[1] It is noteworthy that the basic procedures followed for the conduct of that enumeration have changed relatively little in the succeeding century. The various headings in Table 45 are self-explanatory. Two other definitions of the term 'population' are used in the official population statistics of England and Wales. The population figures published since the end of the last war are 'home', 'civilian' and 'total'. The *home* population includes the members of the Armed Forces of any other nation who may be stationed in this country as well as the indigenous population, but excludes British-born Armed Forces who are stationed overseas. The civilian population comprises the 'home' population less *all* Armed Forces, *i.e.*, British and foreign, in this country, while the *total* population is an estimated figure of the civilian population together with the Armed Forces of England and Wales wherever they may be stationed, *i.e.*, at home or abroad. The 'home' population is the figure most nearly comparable with the pre-war population estimates. The following table taken from the Commentary of the 1955 Annual Review illustrates these figures.

TABLE 46

ESTIMATED POPULATION OF ENGLAND AND WALES, MID-1955, 000's

	Persons	Males	Females
Total	44,623	21,569	23,054
Home	44,441	21,389	23,052
Civilian	43,916	20,879	23,037

Table 47 supplements the data in Table 45 with the inter-census totals for births and marriages and enabling a rough comparison to be made between the figures for each decade by calculating the corresponding rates which are defined in the footnote to the table. The legitimate *fertility* rate is an improved measure of the birth rate since the births are related to the

[1] Evidence of under-counting in 1801, '11, '21 and '31 censuses is provided by the abnormally high inter-censal increases in population. These increases can only be explained by both the natural increase in the population and more complete enumerations as the public became accustomed to the census.

number of married women aged 15-44 instead of the entire population as with the crude rate. The illegitimate rate shown cannot be directly compared with the fertility figures in the adjoining column. It is simply the proportion of all live births

TABLE 47

MARRIAGE AND BIRTH RATES, ENGLAND AND WALES, 1841-1957

Period	Marriages 000's	Marriage Rate[1]	Live Births 000's	Crude birth rate[2]	Legitimate fertility rate[3]	Illegitimate rate[4]	Infant Mortality[5]	Male births per 1,000 female births
1841-50	1,355	16.1	5,489	32.6			153	1,040
1851-60	1,602	16.9	6,472	34.1	281.0	65	154	1,046
1861-70	1,770	16.6	7,500	35.2	287.3	61	154	1,042
1871-80	1,961	16.2	8,589	35.4	295.5	50	149	1,038
1881-90	2,047	14.9	8,890	32.4	274.6	47	142	1,037
1891-1900	2,394	15.6	9,155	29.9	250.3	42	153	1,036
1901-10	2,641	15.5	9,298	27.2	221.6	40	128	1,038
1911-20	3,076	16.6	8,096	21.8	173.5	48	100	1,044
1921-30	3,025	15.5	7,129	18.3	143.6	44	72	1,045
1931-40	3,615	17.7	6,065	14.9	111.1	43	59	1,053
1941-50	3,673	17.2	7,251	16.9	114.0	61	43	1,061
1951	361	16.4	678	15.5	105.4	48	30	1,060
1952	349	15.8	674	15.3	104.5	48	28	1,055
1953	345	15.6	684	15.5	106.3	47	27	1,059
1954	342	15.4	674	15.2	104.8	47	25	1,059
1955	358	16.0	668	15.0	103.7	46	25	1,060
1956	353	15.7	700	15.6	108.2	47	24	1,057
1957	347	15.4	723	15.6	111.5	48	23	1,060

Source Reg General Annual Reviews Part II

[1] Persons married per 1,000 population of all ages
[2] Total live births per 1,000 population of all ages
[3] Number of legitimate births per 1,000 married women aged 15-44.
[4] Number of illegitimate births per 1,000 total live births
[5] Deaths of infants under 1 year of age per 1,000 live births

which is illegitimate. In 1941-50 the illegitimacy rate was 61 per 1,000 births representing that 6.1 per cent of children born in that decade were illegitimate. The relationship between male and female births shown in the final column should be noted. It is a feature of all populations that more males are born than females but the ratio of male to female births is remarkably constant. This fact provides a useful check on the reliability of the registration system in those countries which are only beginning to develop their vital statistics and especially where female babies are considered inferior to males and may not be registered. The infant mortality rate is the number of deaths in the year among live-born babies under one year of age expressed as a rate per 1,000 of all live births recorded in that year. Strictly speaking, this method will give inaccurate results if the number of births fluctuates sharply from year to year, since many of the

deaths in one year relate to births of the previous year. As a rule, however, this conventional method of measuring infant mortality is sufficient for most purposes.

TABLE 48

POPULATION OF ENGLAND AND WALES, 1841–1951

Proportions per 1,000 of each Sex

Sex - Age	1841	1871	1901	1931	1951
Males					
0—	257	261	229	167	169
10—	214	209	209	174	134
20—	172	162	178	174	144
30—	128	125	139	142	149
40—	96	99	105	126	152
50—	64	73	72	110	114
60+	69	71	68	107	138
All Ages	1,000	1,000	1,000	1,000	1,000
Females					
0—	247	248	215	150	149
10—	204	197	197	159	124
20—	184	171	187	169	139
30—	129	130	142	151	143
40—	96	102	105	135	146
50—	65	74	74	112	124
60+	75	78	80	124	175
All Ages	1,000	1,000	1,000	1,000	1,000

Source Reg General Annual Review

Figure 20

POPULATION PYRAMIDS: ENGLAND AND WALES

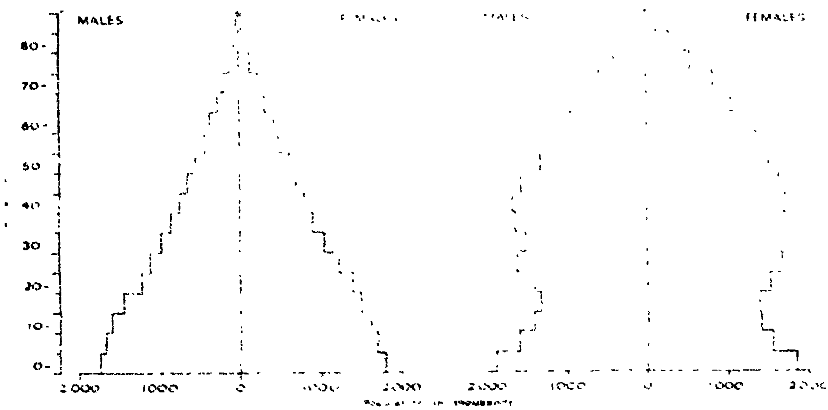


Table 48 gives the breakdown of the population at selected census dates into 10 year age groups for both sexes. The feature of this table is the rising proportion of the population in the higher age groups. This characteristic of the present-day population is brought out by Figure 20. This type of diagram is known as a population 'pyramid'. The 'middle-aged spread' is very marked in the pyramid representing the 1951 population, while the relative youthfulness of the 1891 population reflected in the broad base and narrowing at the upper ages is equally apparent from Figure 20 (page 292).

One reason for this trend is given in the following Table 49 which is based on extracts from a series of life tables which has been compiled from successive censuses. The calculation of a simple life table, known as an abridged table, is explained later. For the

TABLE 49

EXPECTATION OF LIFE, MALES, ENGLAND AND WALES

Age	1841	1870-2	1900-2	1910-2	1930-2	1950-2
0	40·2	40·4	45·9	51·5	58·7	66·4
5	49·6	49·8	54·1	57·1	60·1	64·0
10	47·1	46·7	50·1	53·1	55·8	59·2
15	43·3	42·7	45·7	48·6	51·2	54·4
20	39·9	38·9	41·5	44·2	46·8	49·6
25	36·5	35·4	37·4	40·0	42·5	45·0
35	29·8	28·7	29·5	31·7	33·9	35·6
55	16·7	16·1	15·9	16·9	17·9	18·3
75	6·5	6·0	6·1	6·5	6·4	6·7

Source Reg. General Annual Reviews, Part I

present it is sufficient to note that the above data indicate that a male child born in 1841 would on average have lived only 40·2 years; one born in 1951, however, could expect to live 66·4 years. The female expectation of life at all ages is higher than that for the corresponding male age group but reflects the same trend as for males. The improvement in the expectation of life for both sexes over the past century is directly attributable to the marked reduction in the death rate in the first year of life which is reflected in the rapid and continuous fall in the infant mortality rate (Table 47). The chances of survival in the middle years of life are also better now than in the nineteenth century, but the change is nothing like so marked as in the first year of life. In contrast, as is clear from Table 49 above, there has been little significant improvement in the mortality rates of the highest age groups.

Birth and Fertility Rates

Any statistical analysis at some stage involves the comparison of data relating to different populations or at different periods for the same population. Much has been made of the point in earlier chapters that the first rule of all statistical analysis is to compare 'like with like'. This rule is especially important in the case of vital statistics. In 1939 there were 614,479 live births, 1947 – 886,633 and in 1954 – 673,651. Such annual comparisons are more usually made by calculating what is known as a *crude birth rate* derived by expressing the number of births as a proportion of the population in that year and expressing the fraction per 1,000. Thus for 1939 the rate was 14·8 per 1,000, *i.e.*, $\frac{614,479}{41,642,000} \times 1,000$. The corresponding crude rates for 1947 and 1954 were 20·5 and 15·1 per 1,000. The reader will appreciate that the main fluctuations in the annual crude rates can be explained by the circumstances of those years, but a valid comparison can hardly be made unless it can be assumed that the age and sex structure of the population has remained constant over that period. The so-called crude birth rate is a useful but only approximate measure of a nation's fertility which determines whether or not its population will increase or decrease in the long run. For example, it puzzled many people before 1939 why so much concern was expressed regarding the falling crude birth rate when the total population continued to rise. All that was happening was that people were living longer and the number of deaths for the time being was low. Despite the lower birth rate the effect of lower mortality on the population size was to leave a larger number alive but of greater average age. Ultimately, the population would fall sharply as a large proportion of its members reached old age and died about the same time.

The number of births can be expressed as a rate per 1,000 of the population: this is the crude rate discussed above. Since the population at any time contains a considerable number of children and aged people who are incapable of reproducing themselves, the practice of relating births to such a base is justifiable only on the grounds of its simplicity. The number of children born at any time depends primarily on the number of married women in the population between the ages of 15 and 44,

i.e., the reproductive years. The term 'primarily' is used since births are not dependent solely on the actual number of women in that age group. Since they are nowadays able to regulate the size of their families, the number of births can fluctuate from year to year without any change in the number of such women. Also the number of illegitimate births is relatively small; Table 47 shows that it tends to be around 5 per cent of all births. It follows, therefore, that the key to the birth rate is the number of married women aged 15-44. If the number of live births in a year are expressed as a rate per 1,000 of married women in that age group, we get what is known as the legitimate fertility rate. The Registrar General also calculates an illegitimate fertility rate based on the number of illegitimate births and the number of single and widowed women aged 15-44.

The trend of the legitimate fertility rate during the past century is shown in Table 47. From the same table it appears that while the number of births from decade to decade has not changed very greatly, the growth in the size of the population and the inevitable accompanying increase in the number of married women of reproductive ages has greatly increased. In consequence the fertility rate has fallen steadily throughout the period since 1871.

Unfortunately, even the fertility rate suffers from certain weaknesses as an indication of population trends. To assume that the fertility of women remains unchanged during the entire reproductive period is to ignore all the evidence. Current experience shows that it is the first 7-10 years of married life which are the most important for child-bearing. Thus, in the long run, if the age structure of the married female population changed markedly within the range 15-44 years, the fertility rates would be affected. Such distortion of these rates can be overcome by a technique of age standardisation. This involves comparing the actual number of live births recorded for a female population of a certain age structure with the number of births which would have occurred if females at those ages had borne children at certain selected rates. These 'standard' rates may be selected quite arbitrarily, but rates for 1938 have been used by the Registrar General for post-war comparisons. Various techniques, some of considerable complexity, have been employed by demographers to meet these problems of measuring the trend of

fertility. The best known and one of the simplest of these measures is a reproduction rate, which is simply an index of the extent to which the current female population of reproductive age is replacing itself.

Reproduction Rates

This index of replacement is obtained by estimating the average number of female babies born to women at each age during the reproductive period of their lives, if they were to experience certain fertility rates. The basis of the index is fairly obvious; if each woman in the present generation bears one female child to replace her, then the population will remain constant. Note that the ratio of female to male births is fairly constant so that in the population as a whole, for each girl there is a boy child born to replace the husband. For purposes of calculating this index of replacement the female population is classified into 5 year age groups and usually, but not invariably, current fertility rates are taken as a measure of future fertility. By expressing the expected number of births to all women in any given age group, *i.e.*, married and single, as a rate per thousand women, we derive the *age specific* fertility rate. By aggregating the age specific fertility rates for each five year age group and adjusting this rate for female babies only, we obtain the Gross Reproduction Rate. That is usually written G.R.R. Although both a male G.R.R. and a joint G.R.R. may be calculated, it is customary, and for most purposes sufficient, to calculate the female G.R R. only. This rate has the value of 1 (unity) if sufficient female babies are born to replace the present generation of women of reproductive age. If the G.R.R. is greater

Ages	Total number of women '000's	Total number of births M & F	Age Specific Fertility Rates per 1,000 women
15-19	1,424	27,639	19·41
20-24	1,531	226,817	148·20
25-29	1,653	280,506	169·70
30-34	1,658	194,526	117·30
35-39	1,741	113,966	65·48
40-44	1,669	32,363	19·39
45-49	1,561	2,215	1·42
		878,032	540·90

than unity, then it signifies that current fertility is more than adequate to ' replace ' the present population. If lower than unity, then the reverse applies.

The calculation of the G.R.R. is illustrated on previous page from data relating to 1947.

For the 15–19 years group, the rate of 19·41 per 1,000 is derived by dividing the total number of women in that age group into the number of births, *i.e.*, 27,639 divided by 1,424. The calculation is repeated for each age group.

Since the above data relate to one year's births only, and each age group covers five years, the fertility rates have to be adjusted to the same basis. The aggregate of the various age-specific fertility rates is 540·9 for one year; if the same fertility rates were experienced for five years the figure would be 540·9 × 5 or 2,704 children per 1,000 women.

The total five year fertility rates must next be adjusted to take account of female babies only. Of the 878,032 births occurring in 1947 it is known that 426,024 were female; thus the proportion of female babies is given by $\frac{426{,}024}{878{,}032} = 0{\cdot}485$. The aggregate of all rates for five years equals 2,704 and this is multiplied by the ratio of females in the babies born, *i.e.*, 0·485. This yields a product of 1,312 female children which, expressed per 1,000 women is equal to 1·31 (to 2 dec. places) female children per woman. This is the female G.R.R. which can be described as the ratio of the total number of daughters born to the number of women in the original cohort; in this case 1,000. Sometimes the age specific fertility rates based on female live births alone are given; if this is the case the correction just illustrated is, of course, unnecessary. Then the female fertility rates for each age group are multiplied with the number of women in each age group and the products summed to give the female G.R.R.

The G.R.R. takes no account of the fact that some of this cohort or generation of women will not survive to 15, while others will die at intervals as they pass through the reproductive period. It overestimates, in effect, the number of children this generation of women will bear. If, however, we calculate through the lives of these women the proportion of survivors at each age, and apply only to these survivors the fertility rates corresponding to the successive age groups, we can then estimate the number of

female babies who will be born to replace the present cohort of women as potential mothers. Such a rate is known as the Net Reproduction Rate, or N.R.R.

The calculation of this rate is slightly more complex and is given below. The third column headed l_x shows the number of women in a cohort of 10,000 female live births who survive to given ages, *e.g.*, 15, 25, and so on, until the end of their reproductive period. These figures are derived from the female life table. The construction of such a table is discussed later. The figures in the l_x column signify that of every 10,000 female babies born, 9,645 of them will, if subject to the mortality rates on which the life table is based, survive to their fifteenth birthday, 9,607 to their twentieth and so on.

(1) Age x	(2) Live births per 1,000 women or Age-specific F.R. per 1,000 (f_x)	(3) l_x	(4) $l_x + l_{x+5}$	(5) $\frac{l_x + l_{x+5}}{2}$	(6) $5L_x$	(7) All babies born to generation of 10,000 $5L_x \cdot f_x$
15–19	19·41	9,645	19,252	9,626	48,130	9,342
20–24	148·20	9,607	19,161	9,580	47,900	70,990
25–29	169·7	9,554	19,043	9,522	47,610	80,790
30–34	117·30	9,489	1[illegible],905	9,452	47,265	55,440
35–39	65·48	9,416	18,740	9,370	46,850	30,680
40–44	19·39	9,324	18,525	9,263	46,315	8,982
45–	1·42	9,201	18,206	9,103	45,515	646
		9,005				
	540·9					256,870

The next column (4) is derived by adding together successive values in the l_x column, and the figures in column (5) are the result of dividing the aggregates in the preceding column by 2. The purpose of these calculations is merely to adjust – albeit approximately – the number of women in each age group to the number of women alive mid-way through each of the relevant five year periods. This figure gives an estimate of the number of women in each age group who, on average, in each five year period are alive to bear a child. Since we are dealing with five year groups, the foregoing average must be multiplied (in column 6) by 5 to give the effective number of women in the various age groups capable of reproducing themselves during the

five year period. That column is headed $5L_x$ and is simply the total years of life lived within each five year period by the survivors to those ages of the original 10,000 female babies.

To each figure in the column headed $5L_x$ we apply the corresponding age-specific fertility rate, giving the products in the next column (7). The sum of these products represents the total number of babies of *both* sexes which will be borne by the original 10,000 females by the end of their 44th year if their rate of survival is as shown by the l_x column (column 3) and the fertility rates for each age group are as given in the second column. As before, the total is then corrected for female births only, *i.e.*, 256,870 × 0·485 = 124,582. This total is then expressed as a rate per 1,000 women, *i.e.*, $\frac{124{,}582}{10{,}000} = \frac{-}{1{,}000}$ = 1·245 or 1·25 (to 2 dec. places) which is the value of the Net Reproduction Rate for 1947.

If the N.R.R. is equal to unity, it signifies that despite mortality among the women who may be expected to bear children, the current female population is replacing itself. If the N.R.R. is above unity then the population is being more than replaced. Between 1861–81 the N.R.R. reached a rate of 1·50 which is equivalent to the population doubling itself every thirty years. Such a rate was quite exceptional, it seems that for much of the last century the rate of births in this country was as high as has been recorded anywhere in the world. If, as was the case just before the last war, the N.R.R. is below unity (it was 0·81 in 1935–8), then if it remains below unity at some future date a fall in the population size is inevitable. In the immediate post-war period, the N.R.R. in England and Wales was above unity, *i.e.*, the rate of births was more than sufficient to ensure the maintenance of the then existing population.

When first developed by Dr. Kuczynski before the war, the N.R.R. was believed to provide the basis for reliable population projections, but the limitations of the index, for it is primarily an index of replacements, soon became apparent. The main weaknesses of the N.R.R. arise from the need to make two assumptions which virtually determine the result. The first is the need to assume a rate of female mortality to determine the Net R.R. as distinct from the Gross R.R. Mortality is not liable to significant fluctuations, so that as a potential source of error any forecast

regarding its future course is probably not very serious. In contrast, however, the age specific fertility rates used in calculating the G.R.R. are forecasts of future fertility. It is customary to use the rates current at the time the forecast is being made or some alternative which can perhaps be better justified. But they remain no more than guesses. Future fertility depends primarily on two factors, *i.e.*, the rate at which people marry and the rate at which married couples build up their families, and both of these factors are highly susceptible to change even from year to year. It follows therefore that any assumptions regarding the fertility rates appropriate for the calculation of the N.R.R. must postulate certain behaviour patterns in respect of these two influences. In other words, the N.R.R. can merely tell us what will happen to the population if the assumptions underlying its calculation remain valid. The statistical sub-committee of the Royal Commission on Population considered various alternatives to determine the extent to which a given population was currently replacing itself. Attempts have been made to take into account the duration of marriage and the size of the existing family of married couples in estimating future fertility, but so far the demographers have not succeeded in devising a reliable measure of population trends.[1]

Life Tables

Table 49 showed how the expectation of life had improved over the past century. Those data were based upon successive life tables which have been prepared for the Registrar-General at intervals throughout the past century. The practice has now been established that a new table is prepared on the basis of the data derived from the latest census. The latest official life table is the eleventh which was calculated on the 1951 population using the mean annual death rates for the three year period 1950–52.[2] It is considered better to use the average death rates of such a period rather than the mortality experience of any single year, since that may have been subject to abnormal influences, *e.g.*, an influenza epidemic which markedly increased the death rates for certain age groups.

[1] See, however, the Appendix I to the Report of the Royal Commission on Population.

[2] Life tables are also prepared for inter-censal years (see any recent Annual Abstract of Statistics) but the 'official' English Life Tables are based on the census population as explained above.

Life tables have a long history, the earliest in England being calculated by John Graunt in the 17th century, but it was the development of life assurance which emphasised the need for such tables. The first official or English life table as it is known, was compiled by Dr. William Farr of the G.R.O. in 1841 on the basis of the census returns of that year. These official tables are known as 'English Life Tables' (separate tables are prepared for Scotland) and have been prepared decenially since 1841. The latest is No. 11, an excerpt from which is given in Table 50. These tables are to be found in the slim booklet entitled Life Tables 1951 published as part of the decennial supplement to the Census of Population volumes. Life tables are prepared for each sex and each year of life on the basis of the population census data which provide the most accurate information regarding the age distribution of the population that is ever available.

The life table provides two pieces of information. First, it traces the mortality experience of a hypothetical population

TABLE 50

ENGLISH LIFE TABLE No. 11, 1950-52

(Selected ages only)

Age x	l_x	d_x	p_x	q_x	$\mathring{e}_x$
0	100,000	3,266	96734	·03266	66·42
10	95,866	50	99948	·00052	59·24
20	95,151	12[illegible]	99871	00129	49·64
30	93,820	147	99843	00157	40·27
40	91,968	267	99710	·00290	30·98
50	87,591	745	99150	·00850	22·23
60	75,823	1,796	9[illegible]31	·02369	14·79
70	52,350	2,958	·94349	·05651	9·00
80	21,130	2,880	86371	·13629	4·86

Source: Registrar General's Decennial Supplement, 1951. Life Tables.

from birth to death. More precisely it shows what proportion of a hypothetical population, all of the same age, will survive to any given age. This information is contained in the column headed l_x in Table 50, from which may be read the number of 'lives' surviving from year to year or, in this illustration from decade to decade. Thus, according to the English Life Table No. 11, from a hypothetical 100,000 born, 95,866 males, or nearly 96 per cent, will survive to their tenth birthday, just as their number will

diminish to 91,968 by the fortieth year, if this hypothetical population is subject to the same mortality experience as the actual population on which the table is based. In short, the l_x column is concerned with survivorship from year to year. From these data it is possible to calculate a second set of data which is given in the column headed $\mathring{e}_x$. This is the average expectation of life at any age, i.e. the average future lifetime which would be lived by persons aged x if subject to the death rates on which the life table is based. For example, for a male child at birth the average expectation is 66·42 years, at age 10 the corresponding figure is 59·42 years and so on. It is important to realise that this does not mean that every child born will live 66·4 years; this is nonsense. It simply means that if we add up all the years lived between them by this particular population of 100,000 males, they will live *on average* 66·4 years. Some will live longer, others will die early; but taken as a group they will on average live that span.

In a full life table such as English Life Table No.11 the survivors are given for each year of life from birth to 105. Similarly the expectation of life is given at each single year of age. It is on the basis of such data among other information such as fertility rates that the Registrar-General can estimate the size of the future population. The life offices base their assurance premiums on similar data. Between censuses, when the age distribution of the population can only be estimated, the Registrar-General prepares what are known as *abridged* life tables. These are reproduced in the Annual Abstract of Statistics as well as in the General Register Office publications. Instead of calculating the l_x and e_x values for *every* year of life between birth and ages over 100, he groups the data as in Table 51. The first five years of life are set out singly; then from 5 to 25 the age groups are grouped in five-year intervals, and thenceafter in 10 year groups up to the final groups of 85 and over. This greatly reduces the amount of calculation and while the results are only approximate, as will be seen, the divergence between the results from a full life table and an abridged one is for many purposes not significant. During the ages from 5 to 25 the quinquennial grouping, as with the ten year grouping from 25 to 75, is permissible since the average death rates for each quinquennium or decennium are reliable guides or averages of the death rates for

each of the grouped years. In other words, in such short periods the death rate does not vary significantly. In the first few years, however, especially the first year, the year to year variation is marked and grouping of the first five years would not provide reliable results. At the other extreme, for the age group 85 and over, the mortality rate used is again approximate since it is an average of all the annual rates from 85 to, say, 105. But even if there is here a large degree of approximation, the numbers of survivors is so small that the overall picture is not affected. The calculation of such an abridged life table is illustrated in Table 51.

While the lay-out of Table 51 (pp. 306-8) appears complex, its actual form, as will be seen, is determined primarily by the nature of the calculations. Reference to Table 50 will show that only five columns, apart from the age distribution, are reproduced in the English Life Table No. 11. There are twelve in Table 51 and of them only two, l_x and $\mathring{e}_x$ appear in the English Life Table. All the other columns in Table 51 are working columns. The first seven are needed to provide the l_x values, i.e. the number of lives at each age x, and the remainder to give the $\mathring{e}_x$ values which are derived from the l_x values. The first three columns of Table 51 are self-explanatory. They show the number of males by age enumerated at the 1951 census, adjusted to show the number at mid-1951. The third column of Table 51 gives the total number of male deaths at each age recorded in England and Wales during the three year period 1950-52 inclusive. The next column (3) is the result of dividing the number of deaths at each age by three to adjust them to an annual basis and then by the population of that age given in the second column. The resultant figure is a crude rate of mortality for that age group, written m_x. Thus the rate for the first year of life is ·03402 or 34 per 1,000 of the population of that age. These are known as *central* death rates since they are based on the mid-year population.

For purposes of a life table the crude 'central' mortality rates given above require adjustment. The population figures in column (1) are the 1951 mid-year estimates. Since the deaths occurring in any given year of life are usually spread over the entire twelve months, the actual population of any age which is at risk from the beginning of the year can be estimated by adding half of the annual deaths at each age ($\frac{1}{2}$D) to the mid-year popu-

lation at that age (P). The probability of dying at age x, represented by q_x, can then be written $q_x = \frac{D}{P + \frac{1}{2}D}$. The chances of dying at any age are directly related to the actual death rates which were represented above as $m_x = D/P$. By multiplying both sides by P, this relationship can be re-written $Pm_x = D$. Substituting this in the formula for the probability of dying at age x, *i.e.*, $q_x = \frac{D}{P + \frac{1}{2}D}$, we get $\frac{Pm_x}{P + \frac{1}{2}Pm_x}$ which by eliminating the P's give $\frac{m_x}{1 + \frac{1}{2}m_x}$. This multiplied by 2 to remove the fraction yields: $\frac{2m_x}{2 + m_x}$. This, it may be repeated, is the formula for obtaining the value of q_x, *i.e.*, the probability of dying at age x.

In any year of life x the probability of survival for the period of year x is represented by the symbol p_x just as the chances of dying during that year are represented by q_x. Thus $p_x + q_x = 1$ or certainty. In other words, the chances of survival can be rewritten as $1 - q_x$ and since $q_x = \frac{2m_x}{2 + m_x}$ then $p_x = 1 - \frac{2m_x}{2 + m_x} = \frac{2 + m_x}{2 + m_x} - \frac{2m_x}{2 + m_x}$ which reduces to $p_x = \frac{2 - m_x}{2 + m_x}$. Since the values of m_x for all relevant values of x are given in column (3) of Table 51 the probability of survival (p_x) between any age x to $x + 1$ can be derived.

It is this formula which must be applied to the crude death rate at each successive age. To illustrate the calculation of p_x the numerator and denominator of the fraction $\frac{2 - m_x}{2 + m_x}$ for each value of m_x are set out in column (4) of Table 51.

Take, for example, the first line where m_x is equal to 0·03402, which rounded to four figures is 0·0340. This, subtracted from 2 for the numerator and added to the denominator, gives the values shown in the next column (4), *i.e.*, 1·9660 and 2·0340. In the next column, the logarithms of these values are set out and the one value subtracted from the other. The reader will recall that subtraction of the logarithms of two numbers provides the antilogarithm of the quotient when those two numbers are divided one into the other. The value of the anti-log $\bar{1}$·9853 is 0·9668

which is the value of p_x at age 0. Lack of space makes it impossible to insert a column of p_x values, just as it is not possible to show the q_x values; but they can both easily be derived from the data given and the student may like to perform the necessary calculations. Strictly speaking, the p_x rate for the first year of life needs to be calculated differently, since the average life of those who die in that year is barely 2 months. But this does not alter the basic principles of constructing the life table.

Once the probability of survival during the first year of life is known, we can calculate how many of a generation of 10,000 new born males will survive to their first birthday, *i.e.*, start upon their second year of life.[1] All that is needed is to multiply the number of males alive at the beginning of the year (l_x) by the probability of survival (p_x). Thus, for the first year there are 10,000 males with a probability of survival of 0·9668; so that 9,668 of the original generation of 10,000 embark upon the second year of life. In Table 51 this calculation is shown in column 7; the upper figure is the logarithm of the l_x value, *i.e.*, the log. of 10,000 is 4·0000 and the log. of p_x which was already obtained in column (5) is $\bar{1}$·9853. The anti-log., *i.e.*, their sum, is 3·9853, which is set out below the first two logs and its value 9,668 is inserted in the l_x column (6) indicating that such a number of lives are at risk at the beginning of that next year.

The entire calculation described in the preceding paragraph is repeated for the values opposite age 1. The m_x value is adjusted to give the numerator and denominator of the fraction which gives the value of p_x. The log. of p_x, *i.e.*, $\bar{1}$·9992, is added to the log. of 9,668 shown in column (7) which gives the anti-log., 3·9845, of the product of p_x and l_x where x is *1*. This figure is then carried down to give the number of males alive at the beginning of the third year, *i.e.*, 9,649.

When we come to the age group 5 . there starts a quinquennial as opposed to single year grouping. The fraction for deriving the value p_x has hitherto been based upon single years, but for the five year age groups it is adjusted to $\frac{2 - 5m_x}{2 + 5m_x}$.

This change is explained by the fact that if the mortality in any

[1] It is customary to start with a hypothetical population of male (or female) births. This is often termed a radix. By convention the actual number may be 1,000, or 10,000, or 100,000. The Registrar General's life table proper uses 100,000, while in the abridged table 10,000 is customarily adopted as the original radix. In the above illustration, the hypothetical population will consist of 10,000 males, subject to the mortality rates experienced by males in England and Wales in 1950-52.

TABLE 51

ABRIDGED LIFE TABLE – ENGLISH MALES

Age x	1951 Population (000's) (1)	1950-2 Deaths (2)	m_x (3)	$\frac{2-m_x}{2+m_x}$ (4)	$\log p_x$ (5)	l_x (6)	$\log p_x + \log l_x$ (7)	$l_x + l_{x+1}$ (8)	L_x (9)	T_x (10)	$\log T_x - \log l_x$ (11)	$\mathring{e}_x$ (12)
0	338	34,494	·03402	1·9660 2·0340	0·2936 0·3083	10,000	4·0000 $\bar{1}$·9853	19,668	9,834	664,658	5·8227 4·0000	
					$\bar{1}$·9853						1·8227	66·46
1	353	2,573	·00243	1·9976 2·0024	0·3006 0·3014	9,668	3·9853 $\bar{1}$·9992	19,317	9,658	654,824	5·8161 3·9853	
					$\bar{1}$·9992						1·8308	67·74
2	373	1,592	·00142	1·9986 2·0014	0·3008 0·3013	9,649	3·9845 $\bar{1}$·9995	19,287	9,644	645,166	5·8097 3·9845	
					$\bar{1}$·9995						1·8252	66·86
3	411	1,232	·00100	1·9990 2·0010	0·3009 0·3012	9,638	3·9840 $\bar{1}$·9997	19,270	9,635	635,522	5·8029 3·9840	
					$\bar{1}$·9997						1·8183	65·91
4	429	997	·00077	1·9992 2·0008	0·3009 0·3012	9,632	3·9837 $\bar{1}$·9997	19,257	9,628	625,887	5·7965 3·9837	
					$\bar{1}$·9997						1·8128	64·98

5—	1,616	3,262	·00067	3·9933 4·0067	0·6013 0·6028	9,625	3·9834 $\bar{1}$·9985	19,217	48,042	616,259	5·7898 3·9834	
					$\bar{1}$·9985						1·8064	64·03
10—	1,429	2,335	·00054	3·9946 4·0054	0·6015 0·6026	9,592	3·9819 $\bar{1}$·9989	19,160	47,900	568,217	5·7545 3·9819	
					$\bar{1}$·9989						1·7726	59·24
15—	1,335	3,745	·00094	3·9906 4·0094	0·6011 0·6031	9,568	3·9806 $\bar{1}$·9980	19,091	47,727	520,317	5·7162 3·9808	
					$\bar{1}$·9980						1·7354	54·38
20—	1,427	5,840	·00136	3·9864 4·0136	0·6006 0·6035	9,523	3·9788 $\bar{1}$·9971	18,984	47,460	472,590	5·6744 3·9788	
					$\bar{1}$·9971						1·6956	49·62
25—	3·140	14,969	·00159	1·9841 2·0159	0·2976 0·3045	9,461	3·9759 $\bar{1}$·9931	18,772	93,860	425,130	5·6285 3·9759	
					$\bar{1}$·9931						1·6526	44·93
35—	3,291	28,468	·00288	1·9712 2·0288	0·2947 0·3073	9,311	3·9690 $\bar{1}$·9874	18,355	91,775	331,270	5·5202 3·9690	
					$\bar{1}$·9874						1·5512	35·56
45—	2,955	71,444	·00806	1·9194 2·0806	0·2830 0·3182	9,044	3·9564 $\bar{1}$·9648	17,385	86,925	239,495	5·3793 3·9564	
					$\bar{1}$·9648						1·4229	26·48

TABLE 51 (*continued*)

ABRIDGED LIFE TABLE – ENGLISH MALES

Age x	1951 Population (000's) (1)	1950-2 Deaths (2)	m_x (3)	$\frac{2 - m_x}{2 + m_x}$ (4)	$\log p_x$ (5)	l_x (6)	$\log p_x + \log l_x$ (7)	$l_x + l_{x+1}$ (8)	L_x (9)	T_x (10)	$\log T_x - \log l_x$ (11)	$\mathring{e}_x$ (12)
55–	2,032	140,845	·02310	1·7690 2·2310	0·2477 0·3485 $\bar{1}$·8992	8,341	3·9212 $\bar{1}$·8992	14,954	74,770	152,570	5·1835 3·9212 1·2623	18·29
65–	1,373	227,322	·05519	1·4481 2·5519	0·1607 0·4067 $\bar{1}$·7540	6,613	3·8204 $\bar{1}$·7540	10,366	51,830	77,800	4·8910 3·8204 1·0706	11·77
75–	545	208,487	·1275	0·7250 3·2750	$\bar{1}$·8603 0·5152 $\bar{1}$·3451	3,753	3·5744 $\bar{1}$·[illegible]	4,584	22,920	25,970	4·4145 3·5744 0·8401	6·92
85–	62	53,031	·285	1·0000 0·285	0·0000 $\bar{1}$·4548 0·5652	831	2·9195	–	3,050	3,050[1]	3·4843 2·9195 0·5648	3·67

m_x = central death rate of age group x to $x + 1$

l_x = number of a generation surviving to age x

p_x = proportion of the generation surviving to age $x + 1$ having reached age x.

L_x = total years of life lived by the generation between ages x and $x + 1$.

T_x = total years of life lived by entire generation to age x.

$\mathring{e}_x$ = average length of life of those surviving to age x.

[1] Since $l_x \cdot \mathring{e}_x = T_x$ then $l_{85-} \cdot \mathring{e}_{85-} = 831 \cdot 3{\cdot}67 = T_{85-} = 3050$.

single year is given by m_x, then in a period of 5 years, each assumed to be the same as the other, the mortality will be five times as great in a population starting at the beginning of the first year of the quinquennium. To facilitate the calculation of m_x whereby we derive the logarithm of $5p_x$, the fraction is multiplied by 2 to yield $\frac{4-10m_x}{4+10m_x}$. It is easier to multiply m_x by ten, than by five! The reader can check the figures in columns (3) and (4) accordingly; 4 minus 10 times m_x, *i.e.*, ·00067, yield 4 − ·0067 = 3·993. The remainder of the calculations are as described above. At age 25 there is a further change; the age grouping being altered from five to ten years. The formula for deriving p_x must therefore be changed from its quinquennial form to 10 years, and so it becomes $2-10m_x$ divided by $2+10m_x$. Having obtained from this formula the log. of $10p_x$ it is added to the log. of l_x, *i.e.* l_{25}, to yield the anti-log. of the number of lives at risk at the beginning of the following period. The calculation is repeated for the remaining 10 year age groups.

For the age group '85 and over' the adjustment to the m_x rate to give the probability of survival (p_x) is different from that used hitherto. According to the full life table prepared by the Registrar General, there are survivors aged 104 years. But the number of survivors between ages of 85 and 105 is so small that detailed calculation is pointless: furthermore the assumption on which the earlier p_x rates were based, that deaths are spread evenly throughout the period is no longer valid. The full calculations for these age groups are far too complex for this book but a good approximation to the value of $\mathring{e}_x$ derived by the Government Actuary is obtained by calculating $\mathring{e}_{85-}$ directly from the formula $1/m_{85-}$. Note that the death rate represented by the symbol m_{85-} is simply an approximation to the average of the death rates at all ages of 85 and over. Hence we write m_{85-} instead of m_{85}, which signifies the death rate of males during their 85th year only. From the data we get 1/0·285 which equals 3·67 years. This corresponds relatively closely with the expectation of 3·48 years in the official 11th English Life Table.

At this stage the information provided by the table enables us to derive the values of p_x and q_x, *i.e.*, the probabilities of survival and death respectively at certain ages. The column l_x shows

the effect of such rates on a hypothetical generation of 10,000 new-born males, as they grow up. Thus 332 (d_x) fail to survive the first year, leaving only 9,668 (l_x) to celebrate their first birthday. The number surviving to their 15th birthday (l_{15}) is 9,568, and at 55 there are 8,341 of the original 10,000 still alive. But this is only part of the complete life tables: *i.e.*, that part comprising the first four columns of Table 50 above, which was taken from the published official English Life Table No. 11. The final column thereof is headed $\mathring{e}_x$, *i.e.*, the expectation of life at age x and it is to the calculation of this figure that we now turn.

Column (8) is headed ($l_x + l_{x+1}$) and the first figure in it, 19,668, is obtained by adding together the first two values in the l_x column (6), *i.e.*, 10,000 and 9,668. The next figure of 19,317 is the sum of the second and third values in the l_x column, *i.e.*, 9,668 and 9,649. For the next value, the fourth will be added to the third and so on all the way down the column. Having done this the totals in column (8) are divided by two to give the values in the next column headed L_x. The purpose of this averaging technique must be clearly understood. Take, for example, the first two values in the l_x column, *i.e.*, 10,000 and 9,668. From these figures it emerges that 332 lives were lost in the first year. Since it may be assumed that the deaths were spread out over the year, the *average* live population under one year of age was 10,000 less 166 ($\frac{332}{2}$) which corresponds to the first value in the L_x column. In other words, the total period of life lived by the original 10,000 males during that first twelvemonth was approximately 9,834 years.[1] At the beginning of the next year there were 9,668 lives at risk, but only 9,649 survived and following the same principle as before, we estimate that they lived in the aggregate 9,658 years between year x and $x+1$.

When we come to the 5 year age groups an adjustment to the figures in column (8) is required. This column headed $l_x + l_{x+1}$ should now be described as $l_x + l_{x+5}$, for we are estimating the total years lived by the survivors during the quinquennium and not merely one year as for the age groups under 5. As before, the sum of the two successive values of $l_x + l_{x+5}$ is divided by two and then the quotient multiplied by five to give the value to be entered in the L_x column, which from this point on

[1] Alternatively, if 9,668 survived the first year, they lived in total 9,668 years. 332 died, some at the beginning so that they may have lived only a few days, while others died at the end after 11 months, of life. On average, we may assume the 332 non-survivors lived 6 months apiece, or 166 years. Thus the original 10,000 babies lived in all 9,668 + 166 years, *i.e.*, 9,834.

should be described as $5L_x$. The explanation can be illustrated by reference to the figures corresponding to age 5 —. The $l_x + l_{x+5}$ values are 9,625 and 9,592; their sum is 19.217 which divided by two give the 'annual' population and this when multiplied by five yields 48,042. This is simply the number of years lived by the 9,625 males who entered their fifth year until they reached their tenth birthday or died. This can be checked by a simple calculation. If the entire group of 9,625 males who entered upon their fifth year had survived until their 10th birthday, they would have lived altogether 5 × 9.625 years, *i.e.*, 48,125. In fact 33 of them (9,625 – 9,592) died before the 10th birthday and if we assume they lived on average 2½ years, they lost by their premature deaths 2½ × 33 years of life, or 83 years in all. This figure taken from 48,125 equals 48,042 years, which value is to be found in column (9). These calculations are repeated for all the quinquennial age groups to 20—.

The next age group 25— is a ten year grouping so that after adding successive values in column (8), dividing by two and then multiplying by five, instead we now multiply the quotient by *ten* for the reasons already stated and get the values in column (9). In this case they are $10L_x$.

When we reach the final age group there is only a solitary l_x value left, in this case 831. The L_x value can only be derived indirectly from the $\mathring{e}_x$ value of 3·67, the derivation of which was explained above. The values of T_x in column (10) are obtained by aggregating all the L_x, $5L_x$ and $10L_x$ values. Without L_{86} the T_x column could not be obtained. Since e_x is obtained by dividing T_x by l_x, it is possible to calculate the value of T_{86-} from the values of $\mathring{e}_{85-}$ and l_{85-} which are already known. Thus we get 3·67 × 831 = 3,050 to give the value of T_{85-} which corresponds in the final age group, as will be seen below, with the figure for L_{86}.

By adding up all the values in the L_x column we get a total of 664,658 years. This is the total duration of the lives of the entire 10,000 males from birth until death. Since we started with 10,000, the average length of life was 664,658 ÷ 10,000, *i.e.*, 66·46 years which is the first figure shown in the final column headed $\mathring{e}_x$. To obtain the expectation of those who reach their first birthday, we again add up all the years lived by the 9,668 males who started from age 1 and divide the total of 654,824 by 9,668 so that the average expectation of life for those surviving

their first year is 67·6 years.[1] The student may now perhaps understand why it is possible for the expectation of life to increase after having lived one year, *i.e.*, at birth it is 66·4 and at age 1 it is 67·6! The reason is that both values are averages and in calculating the second, the extreme low values (*i.e.*, average life 6 months) of which there are a large number (332) are dropped. The student will recall the discussion of the arithmetic mean and how it was affected by extreme values, particularly where the corresponding frequencies were high.

It is interesting to compare some of the values for $\mathring{e}_x$ in Table 51 with those shown in Table 50 which is taken from the official life table No. 11. It will be seen that the divergence at any age is at most 0·1 of a year. Thus, it can be seen that even with simple approximate calculations, results similar to those arrived by using what are known as graduation formulae can be obtained.

The value of life tables is not solely that they are used for life assurance, or for making population projections, important as these undoubtedly are. Life tables can also be used, although they are rather cumbersome for this purpose, for comparing the mortality experience of different populations, i.e. different in location as well as different in time. Thus a country whose population has a high mortality in the early years of life will find this fact reflected in a more rapidly diminishing l_x column than a country with a lower death rate. Similarly, the $\mathring{e}_x$ values at each age will be lower. It is more convenient, however, to compare relative mortality (and fertility) in different populations by means of what are known as *standardised* rates. These are explained below. One last application of the life table principle may be mentioned. Instead of measuring the loss each year in a population from death, we can show the 'loss' in a generation of spinsters to marriage. Starting at age 16 with a radix of 10,000 spinsters, we can calculate the 'probability of marriage' instead of death, using the marriage rates of spinsters in a given period as the m_x values. Both a gross and net nuptiality table can be prepared, the former taking only marriage into account, while the *net* table also allows for deaths among spinsters before they marry. An example of a net nuptiality

[1] In practice, the same result is achieved by taking away from the aggregate the total lives lived by the previous age-group, *e.g.*, 664,658 less 9,834 = 654,824.

table is given in the Commentary volume for 1956.

REFERENCES

An interesting application of life table principles is to be found in 'The Length of Working Life of Males in Great Britain', H.M.S.O. 1959. There is also a summary account of the method of constructing a life table which will supplement the above account. The decennial supplement on the E.L.T. No.11 also gives an explanation of the construction of the table, but the exposition is mathematical since a different technique is used.

Standardised Rates

For 1957 the Registrar General's Annual Review (Part I, Table 12) gives the following information:

Town	Population	Crude Death Rate	Crude Birth Rate
Clacton ..	24,890	15·4	9·8
Stevenage ..	26,000	5·3	31·2

According to these figures the death rate in Clacton is almost three times as high as in Stevenage. One possible interpretation is that the former seaside resort is a thoroughly unhealthy place! This view could even be supported by reference to the birth rates which reveal a three to one ratio in favour of Stevenage. From this we might also conclude that the air in Stevenage is rather more bracing than the much vaunted ozone of the east coast! Both these inferences are of course nonsensical in the light of our knowledge of these two districts. On the other hand, what is the difference in respect of births and mortality in the two towns if it is not as shown in these crude rates?[1]

According to the first chapter of this book the primary object of compiling statistics is to enable comparisons to be made. It is essential to be able to compare the birth and death rates between different areas of the country, but the contrasting figures quoted above suggest that the crude rates are unsuitable for comparative purposes. For a short period comparison of mortality and fertility experience in the population of *one* area the crude rate may be used since changes in the sex and age structure therein are *usually* slow to emerge. Differences in such rates, *e.g.*, the crude death rate between two towns or regions, however, can be

[1] The relevant figures are given on p. 321. In passing it may be mentioned that the New Towns, of which Stevenage is one, consist largely of young couples and the birth rate is abnormally high – an average of about 30 per 1,000, or double the rate for England and Wales as a whole

attributed either to local conditions affecting mortality, a high proportion of aged in the population, or a combination of both. There are almost invariably differences in the age and sex composition of the populations in different areas and the first step to comparing mortality (or fertility) experience in several areas is to eliminate or compensate for the influence of these factors. This is done by calculating in place of the crude rate, what is known as a *standardised* rate for each area's population.

The basis of standardisation as it is known is illustrated by a simple hypothetical example below. Two towns, A and B, with populations of about 80,000, record crude death rates of 14·5 and 13·7 per 1,000. The derivation of these figures will be apparent from the three columns under each main heading. For example, in Town A, in the age group 0—, there were 40 deaths in a population of 5,000 under 5 years which equals a death rate of 8 per 1,000. The total figures, *i.e.*, deaths ÷ population, give the crude rates. Note that the crude rates for the two towns are not derived by adding the age specific death rates. The crude rates are in effect the weighted arithmetic means of the age specific death rates for A and B respectively. It is apparent that although the difference between the crude rates is relatively slight, the age specific death rates in Town B are higher for each age group except the last two. To permit a comparison of the mortality experience of the two towns, it is necessary to find out what would happen to a given, termed a 'standard', population in both towns. This population is given in column 8 and the next two columns show the number of deaths to be expected if the standard population were subject to the age specific death rates of towns A and B. Expressing the expected deaths as a rate per thousand of the standard population, the standardised death rates for A and B are 8·9 and 10·6 per 1,000 respectively, *i.e.*, $\frac{890}{100{,}000}$, etc. In other words, the death rates are higher in B than in A and the opposite conclusion, which could be drawn from a comparison of the crude rates, was due entirely to the more youthful population of B. This method of standardisation is known as the *direct* method.

It will probably have occurred to the reader that with this method of standardisation the actual 'standard' population employed is a matter of indifference. Any population would do,

TABLE 52

CALCULATION OF STANDARDISED DEATH RATES (DIRECT METHOD)

Age Group	Town A			Town B			Standard Population 000's	No. of expected deaths in standard population if subject to mortality of	
	Population 000's	No. of Deaths	Death Rate per 1000	Population 000's	No. of Deaths	Death Rate per 1000		A	B
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
0—	5	40	8	7	70	10	12	96	120
5—	15	15	1	10	20	2	24	24	48
15—	15	15	1	15	60	4	20	20	80
25—	15	30	2	19	114	6	30	60	180
45—	25	375	15	25	500	20	10	150	200
65—	5	400	80	3·5	210	60	3	240	180
85—	1	300	300	0·5	125	250	1	300	250
	81	1,175	14·5	80	1,099	13·7	100	890	1,058

always provided that a straightforward comparison between the mortality of different years or regions is all that is required. Unfortunately, if the standard population were to be abnormally young, or old, then the resultant rates could be misleading. The point is that demographers still use the crude rates as a first approximation; a death rate of between 10 and 15 per 1,000 is typical of a progressive modern economy while a rate of 30-35 per 1,000 indicates an under-developed community with a high death rate. It is desirable therefore that the standardised rates should approximate to this range of crude rates and for this reason the standard population used for calculating standardised rates should reflect the age and sex structure of present-day population.

From the beginning of this century the Registrar General used as the 'standard' population the 1901 census population, but by the time the second World War started, the 1901 population had ceased to be representative. It was an abnormally 'young' population while the population of the 1930's was markedly more middle-aged. Furthermore, the standard was remote in time; a more up-to-date standard was desirable. The Registrar General might well have taken the census population for 1931 as his new standard, but this too would in time have become out-dated and in any case, with the introduction of a new standard all the standardised rates for the earlier years would have needed recalculation. There would also be a consequent distortion of those

rates for the earlier years since the new 'standard' was so different from the old.

A new method was introduced in 1941 which partly overcomes the problem of an out-dated standard population, by using a form of moving base. The standard population is derived by averaging the population of 1938 and the other year under review. The hypothetical number of deaths which would occur in this standard population if it were subject to the mortality of (*a*) 1938 and (*b*) the year under review, is worked out for both years and expressed as a rate per 1,000. The two standardised rates are then expressed as a ratio, obtained by dividing the 1938 rate into the other. This ratio is called the *Comparative Mortality Index* or the C.M.I. Separate indices are calculated for persons, males and females. The base year is 1938; for which year the C.M.I. has a value of 1·000, *i.e.*, unity. Note that the two standardised death rates for 1938 and the given year are derived in exactly the same way as was illustrated in the example of Towns A and B above. The only difference is that instead of using fixed 'standard' population, the present method uses the mean or 'intermediate' population between 1938 and the given year. Given the two standardised rates, *e.g.*, for the years 1938 and 1955, the 1938 rate is divided into that for 1955 to give the C.M.I. for 1955. The C.M.I. for persons in 1955 was 0·805, denoting that age for age mortality rates were almost 20 per cent lower than in 1938.[1] Note that the C.M.I. permits comparisons to be made only as between the rates in different years for the same category of individuals, *e.g.* males or females, or persons. The male C.M.I. for 1955 may be compared with the same rate for males in any other year. It may not be compared with the female or persons C.M.I. for that year or any other year.

Comparative Mortality Indices measure, in relation to the base year 1938, the trend of mortality over a period of years. Indices have been worked out for each year back to the beginning of the century as well as for each quinquennium in the period 1841–1951 (Table 3, Medical vol.). Reference to Table 4 of any recent Medical volume of the Registrar General's Review will reveal a column headed 'Mortality Ratio'. This is derived for each year by dividing the C.M.I. of that year by the C.M.I. of the previous year. Thus, the C.M.I.'s for persons in 1954 and

[1] For a full explanation of the calculation of the C.M.I. see Reg. General's Review, 1940-5, 1941, Part I, pp. 6-11, also Review for 1941, Part I. Appendix.

1955 were 0·789 and 0·805 respectively, so the Mortality Ratio for 1955 is therefore 1·020, *i.e.*, 0·805 divided by 0·789, signifying that male mortality rose in 1955 from the level of the previous year by 1·6 per cent. As with the C.M.I., so the Mortality Ratios have been worked right back to 1841-50.[1] As from 1959, however, the C.M.I. will be replaced by the Standardised Mortality Ratio. This is discussed below (p. 321).

The direct method of calculating standardised rates for different populations depends on there being available the number of deaths in each population classified by age. For many local, *i.e.* urban or regional populations, these data are not always readily available and it is therefore impossible to calculate local age specific death rates which are then applied to a standard population. On the other hand, the age specific death rates for England and Wales are known and it is often more convenient to apply such standard death rates to the local populations. While the age specific death rates now used by the Registrar General as the standard rates are those of the census year 1951, between 1949 and 1953 the mean death rates of 1947 and 1948 were used. When this method was introduced in 1934, the standard was the average death rates for 1930-2. These continued to be used up to the war but the publication of the series was then suspended until 1949 owing to the large shifts that were taking place in the population.[2]

In principle this method of standardisation is just the same, for we are measuring the mortality experience of two or more populations on a common standard. This method, *i.e.* standard death rates to local populations, is known as the *indirect* method. But unlike the direct method, described earlier, it does not immediately produce a standardised rate. At the first stage of the calculation if merely shows the number of deaths that each population would sustain if they experienced identical, *i.e.* the standard, age-specific death rates. In other words, the first stage of the calculation is to provide a means of adjusting the crude local death rates in respect of any difference between them that is attributable to the fact of the differing age structures of their populations. This correction factor is known as the *Area Com-*

[1] See Tables 3 and 4 of any recent Medical volume of the Review.

[2] See Registrar-General's Review 1954, Part III, pp. 30 and 57. Note that the A.C.F. was introduced in 1934 for deaths only; it was not used for standardising birth rates until 1949. Note, however, that this method of comparing fertility is subject to error since no adjustment is made for the differing proportions of women married in each population.

parability Factor, usually abbreviated to A.C.F. These factors are calculated for the population of each district. They are used for calculating standardised rates whereby the mortality of different *areas* in a given period can be compared. Note that the C.M.I. was designed to reflect changes in mortality (or fertility) in a given population over time. The A.C.F. is used for geographical comparisons. It is not, as will be seen below, itself a death rate as such; it is a standardising factor which when applied to local crude death rates enables the resultant rates for various areas to be compared.

TABLE 53

CALCULATION OF AREA COMPARABILITY FACTORS FOR EASTERN REGION AND TYNESIDE CONURBATION, 1954

Age Group	Est. Mid-Year Popln. 1954. 000's		Number of Deaths 1954. 000's		Standard D/R per 1,000 England & Wales 1951	Expected Deaths on basis of 1951 death rates	
	Eastern Region	Tyneside Con-urbation	Eastern Region	Tyneside Con-urbation		Eastern Region	Tyneside Con-urbation
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0—	248	70	1,279	492	6·53	1,619	457
5—	491	128	189	61	0·51	250	65
15—	419	107	306	85	0·95	398	102
25—	464	131	511	197	1·45	673	190
35—	448	113	924	315	2·63	1,178	297
45—	445	117	2,367	851	6.90	2,990	807
55—	342	87	4,587	1,551	18·0	6,156	1,566
65—	255	58	9,234	2,622	46·1	11,755	2,674
75—	146	27	15,916	3,492	138·0	20,148	3,726
	3,258	838	35,313	9,666	12·5	45,167	9,884

	Eastern Region	Tyneside Conurbation
Index death rate per 1,000 ..	$\frac{45{,}167}{3{,}258} = 13{\cdot}86$	$\frac{9{,}884}{838} = 11{\cdot}8$
Crude rates, per 1,000	$\frac{35{,}313}{3{,}258} = 10{\cdot}8$	$\frac{9{,}666}{838} = 11{\cdot}5$
A.C.F. = $\frac{\text{Standard death rate}}{\text{local index death rate}}$	$\frac{12{\cdot}5}{13{\cdot}86} = {\cdot}90$	$\frac{12{\cdot}5}{11{\cdot}8} = 1{\cdot}06$
A.C.F. × local crude death rate = local adjusted death rate ..	${\cdot}90 \times 10{\cdot}8 = 9{\cdot}72$	$1{\cdot}06 \times 11{\cdot}5 = 12{\cdot}19$
Ratio of $\frac{\text{local adjusted death rate}}{\text{standard rate}}$	$\frac{9{\cdot}72}{12{\cdot}5} = 0{\cdot}78$	$\frac{12{\cdot}19}{12{\cdot}5} = 0{\cdot}98$

In Table 53 the basis of the calculation of the A.C.F. for two areas is illustrated. They are the Eastern region which is one of the standard regions and the Tyneside conurbation, for both of which the Registrar General publishes separate data. The first three columns give the age distribution of the two populations, while the next two columns (4 and 5) show the actual number of deaths recorded for each age group in each region. The total of these deaths expressed as a rate per thousand of the population of each area would yield the *crude* death rate. As the 'standard' the Registrar General now uses the 1951 death rates for England and Wales. These are given in column 6 and are applied to the two populations in turn. The final columns (7 and 8) show the number of hypothetical deaths which would have occurred in the two areas had they both experienced the same mortality as England and Wales in 1951. By expressing the hypothetical number of deaths for each region over its population we get what are known as *index* death rates.

Note that these index rates do not provide any measure of the mortality experience of the two populations. They are merely indices which reflect the relative age structures of the populations of Tyneside and the Eastern region. Since the former's rate is below that of the latter (11·8 to 13·9) and each age group in the two populations has been subject to the same mortality, it follows that the population of Tyneside has relatively fewer persons in the age groups with high mortality rates, *i.e.*, the very old and the very young. The calculation of actual standardised rates for Tyneside and the Eastern Region is as follows: First the two index rates are divided into the standard rate, in this case the crude death rate for all ages in England and Wales for 1951 (column 6). This process yields what is known as the Area Comparability Factor. This is a *standardising* factor which when multiplied with the crude death rates for each region yields the standardised rates for those two regions. These can then be compared.

The purpose of the standardising factor is as follows: The index death rate, it will be recalled, served to indicate the relative favourableness of the age distribution of the two populations in regard to mortality. We derive, by dividing them into the national or standard death rate, a common measure of the extent to which the two populations are either favourably or unfavourably placed. Thus, if the index death rate for one of the

populations is lower than the standard, as in the case of Tyneside, 11·8 compared with 12·5, it signifies that on the whole its age structure is more favourable, *i.e.*, younger than that of England and Wales. It ought, therefore, other things being equal, to have a lower death rate. When this index rate is divided into the standard rate the resultant standardising factor will be greater than unity. When this, in turn, is multiplied with the crude rate for the same population the resultant standardised rate will be higher than the crude rate. The process can be illustrated by reference to the figures from Table 53. The A.C.F. or standardising factor for Tyneside is 12·5/11·8 = 1·06. The crude rate is 11·5 per 1,000 which when multiplied by the A.C.F. gives a standardised rate of 12·19. The standardised rate is higher than the crude rate because the 'favourable' age structure of the Tyneside population yields a low crude rate. The latter when adjusted to take account of this advantage is higher. With rhe Easter Region's unfavourable age structure the position is reversed, *i.e.*, the A.C.F. is less than unity and the standardised tate is therefore lower than the crude rate.

Reference has been made to the Mortality Ratio which the Registrar General calculates to permit a direct comparison of the C.M.I.'s of successive years to be made with each other, instead of merely comparing each year against a single base year, at present 1938. With the A.C.F. a similar ratio is calculated and published with the A.C.F. and crude rate. This is known as the *ratio of the local adjusted death rate to the standard rate*. As may be seen from the last line of calculation following Table 53, it is the standardised local rate divided by the national death rate. This particular ratio for any region can be compared directly with the same ratio for any other area in that particular year.[1] The ratio for England and Wales is unity.

Standardised indices for both births and deaths are published in the latest Annual Reviews of the Registrar General.[2] For comparisons over *periods* one table shows for both births and deaths the crude rates, the C.M.I. and the Mortality Ratio for individual years back to the beginning of this century. Similarly, for the *regions* and administrative areas of England and Wales

[1] An attempt is also made by the calculation of what is known as the *Time Comparability Factor* to adjust the A.C.F. so that changes in regional death rates can be compared over time. This is not a satisfactory method when the population age structure is changing.

[2] Part I, Table 12. Part II, Table E.

the tables give for both births and deaths the crude rates, the A.C.F. and the ratio of local adjusted rate to the standard rate. Using these data it is found that the A.C.F. for Clacton and Stevenage is 1·23 and 0·69 respectively. Thus the standardised birth rates are: Clacton $9{\cdot}8 \times 1{\cdot}23 = 12{\cdot}05$ per 1,000; Stevenage $31{\cdot}2 \times 0{\cdot}69 = 21{\cdot}7$ per 1,000. The standardised death rates are derived from the crude rates and C.M.I. The C.M.I. for Clacton is 0·72, that for Stevenage 2·05, so that the standardised death rates are 10·8 and 11·1 per 1,000 respectively.

Standardised Mortality Ratio

In the Census report on Occupational Mortality the mortality experience of each occupational group was summarised by means of a Standardised Mortality Ratio (S.M.R.). This is defined as the number of deaths registered of men with a given occupational group at ages 20-64, expressed as a percentage of the number that would have occurred if the death rates in each separate age group within the occupation had been the same as in a standard population consisting of all males in England and Wales. The method of calculating an S.M.R. (all causes of death) is illustrated in the following example for the occupational grouping known as Farmers and Farm managers taken from the 1951 report.[1] Columns 1 and 2 give the age distribution of this occupational group.

TABLE 54

CALCULATION OF STANDARDISED MORTALITY RATIO

Ages (1)	Census Population (2)	Standard death rates per million 1949–53 (3)	Expected deaths in occupation $\frac{(2) \times (3)}{1{,}000{,}000}$ (4)
20—	7,989	1,383	55
25—	37,030	1,594	295
35—	60,838	2,868	872
45—	68,087	8,212	2,796
55—64	55,565	22,953	6,377
Total standard deaths 20–64			10,395

Total registered deaths of farmers and farm managers aged 20–64 = 7,320.

$$\text{S.M.R.} = \frac{7{,}320 \times 100}{10{,}395} = 70 \text{ per cent.}$$

[1] Registrar General's Decennial Supplement 1951, Occupational Mortality, Part II, vol. I, p. 17.

The basis of the calculation of the S.M.R. is the expected deaths in the last column. The standard death rates are based on the registered deaths in the five year period 1949–53 inclusive, the rates shown being annual averages. The population in each age group is at risk for a period of five years so that the expected deaths will be five times the annual rate times the population at risk. On this basis the expected number of deaths will be 10,395, *i.e.*, this is the number of deaths to be expected if farmers had experienced the same mortality as all males between 20–64 in the quinquennium 1949–53. In fact the number of deaths of farmers registered in the same quinquennium was only 7,320. This number expressed as a percentage of the expected number of deaths yields a percentage of 70. In other words farmers and farm managers experienced a considerably lower mortality, age for age, than the male population as a whole.

Standardised Mortality Ratios are calculated not only for 'all causes' of death but also for individual diseases such as tuberculosis and cancer, as well as various types of accident. The mortality rates for each disease are analysed by social class and by sex, married women being distinguished from single women. In addition, the S.M.R. is used to compare regional differences in respect of mortality as well as comparing class and regional experience in respect of infant mortality and still births. As stated earlier, from 1959 the Registrar General intends to discontinue the calculation of the C.M.I. and replace it with the S.M.R. Thus for any year the S.M.R. of any population will be calculated as follows. Standard age specific death rates, presumably those of England and Wales for both sexes, in addition to the 'persons' rate, will be applied to the local population. This will give the 'expected' number of deaths which will then be compared with the actual number of registered deaths in that population. The actual will be expressed as a percentage of the expected number to give the S.M.R. for that population in any given year.

It is inevitable that differing results – in arithmetic terms – will be obtained according to the method of standardisation employed for comparative purposes. But, since the object is comparison of two or more populations either by region or over time, this does not matter greatly. The relative differences in their mortality and/or fertility experience will be shown to be

about the same whatever method is used. In practice, of course, one needs merely to consult the appropriate part of the Registrar General's Review to obtain the figures.

Conclusions

With the possible exception of the Board of Trade, the Registrar General's Office is responsible for the preparation of more statistical data than any other single government department. Although vital statistics may appear to be of limited application, in fact the data available serve as the basis of both central and local administration, forward planning for schools, housing, pensions, etc. The same data serve as checks on the samples used by market research organisations in their study of consumer trends. The Census of Population is the most important statistical event of the decade and its reports of the greatest interest and value. It is because reference is so often made to these published statistics that this lengthy chapter has been included in this book. Actually the methods of calculation employed for preparing these data are well beyond the level of arithmetic needed for the rest of this text. But, as has been stated so often, unless the source and basis of published statistical data are understood, there is always the danger of mis-quoting them. The sole purpose of the foregoing sections has been to provide a simple, even if at times incomplete, survey of the basis of these important statistics. Provided this much is clearly understood, the reader should be safe from the worst dangers.

For the reader who may be especially interested in this branch of statistics some suggestions for further reading, apart from the footnote references given earlier, are given below.

REFERENCES

Vital Statistics, B. Benjamin. Allen and Unwin, 1959.

Reports and Selected Papers of the Statistics Committee. Volume II of the Papers of the Royal Commission on Population. H.M.S.O. 1950.

The Census 1951. General Report. H.M.S.O.

*Demography, P. R. Cox. Institute of Actuaries, 1955.

*Introduction to Demography, M. R. Spiegelman. Society of Actuaries, U.S.A. 1955.

* Both of these texts contain a large amount of descriptive material, but the reader who is unfamiliar with the conventional mathematical notation will find some of the discussion of rates and life tables – which is designed for student actuaries – difficult to follow.

CHAPTER XVI

THE CONSTRUCTION OF INDEX NUMBERS

Every reader of this book will know that in each year since the beginning of the last war, prices of the goods we buy have risen in greater or lesser degree. The price of a loaf of bread has risen from 4d. to 1s.; a hundredweight of coal from 2s 6d. to 10s.; a ready-made suit from £3 to £10. These price changes give an indication of the extent to which prices have risen, but generally speaking it is more convenient to indicate the fall in the real value of money by a single measure. For example, it might be said that since 1939 the cost of consumer goods as a group has risen by 200 per cent., *i.e.*, prices are now three times what they were then. This is a useful device because the degree of change in the prices of various goods differs; for example, bread as quoted above had risen by 200 per cent., coal and the ready-made suit by 300 and 333 per cent. respectively. It is much more convenient if all these changes can be expressed by a single figure. The figure used for comparing changes as between different points of time in what is sometimes referred to as 'the price level' is described as an *index* number of prices. In view of what has been said it follows that the term 'price level' is a misnomer; in any period of time individual prices move differently, but the *average* movement can be calculated and it is this average which is measured by the index number. In fact, an index number is really no more than an average, with both its advantages and shortcomings. The most frequent use of index numbers is for measuring the change over selected periods in *prices* of goods, commodities, assets such as securities and so on. For example, there is an Index of Retail Prices, an Index of Wholesale Prices, and one for prices of stocks and shares quoted on the Stock Exchange.[1]

The foregoing sentence should indicate to the reader that price indices are specially prepared to measure changes in particular groups of prices. There is no index suitable for measuring the

[1] See Chapter XVII.

'general level of prices', but there is an index for measuring the cost of goods and services which the average family in the United Kingdom buys as part of its mode of life. This is often termed a 'cost of living' index. Index numbers are specially designed. For example, changes in the volume of industrial output are measured each month by the index of industrial production; changes in the prices of exports and imports from and into the United Kingdom by indices of import and export prices. All these various index numbers are described in later chapters.

Measuring Price Changes

All prices do not change to the same extent over a given period of time. Some prices rise (or fall) more than others, *i.e.*, they move relatively to one another. The difficulty arises when these relative changes have to be 'averaged'. It is possible to get different answers to what appears to be the same question according to the method used for measuring price changes. Suppose we have four commodities which in 1949 and 1959 cost per pound weight as follows: A 5s.–7s. 6d.; B 10s.–12s. 6d; C 15s.–£1; D £1–£3. In the earlier year the four items could have been purchased for £2 10s. and by 1939 they cost £5. The proportionate increase in the total cost of these goods is given by expressing the actual change in cost over the total cost in the earlier year, *i.e.* £2 10s. divided by £2 10s. which is equal to unity. It is customary to express such changes in percentage terms, so that if the 1949 expenditure is termed the 'base', or 100, then the corresponding value for 1959 is 200, or a 100 per cent. increase. Such an index, which is no more than the percentage change in the *aggregate* expenditure on a collection of goods at different points of time, is sometimes referred to as a *simple* aggregative type of index. As will be seen below, it does not really justify the title of an index.

Now instead of taking the actual price of each commodity, calculate the percentage increase in its cost. A is then 50, B 25, C $33\frac{1}{3}$, and D 200 per cent. higher. If these percentages are added together they total $308\frac{1}{3}$, which apportioned over the four items represents an increase of approximately 77 per cent in the prices. This, it will be noted, differs from the figure of 100 per cent. derived by using simple aggregate expenditures. Generally speaking, the simple aggregative type of index derived by

relating two aggregate expenditures is not used in practice because the index may be distorted by any single *large* absolute change which swamps all the other movements. This, it will be recalled, is much the same as saying that the arithmetic mean is unduly affected by extreme values and may sometimes be unsatisfactory as a measure of a given distribution.

It is nevertheless quite possible to use the actual prices of the goods purchased to calculate a price index, without converting them into percentages. To make such an *aggregative* index, as it is termed, the various prices have first to be multiplied by the physical quantities purchased of each commodity. The products of the prices and quantities are summed and the two totals expressed as ratios of one another.

TABLE 55

CALCULATION OF AGGREGATIVE TYPE INDEX

Commodity	Prices		Quantity	Product of Quantity × Price	
	1949	1959	lbs.	1949	1959
	s.	*s. d.*		*s.*	*s.*
A	5	7 6	4	20	30
B	10	12 6	6	60	75
C	15	20 0	3	45	60
D	20	60 0	1	20	60
	50s.	100s. 0d.	14	145s.	225s.

The calculations can be followed in the above table. In this aggregative index, the average increase in the prices of the commodities is derived by dividing the product of the 1949 quantities and prices into the corresponding product for 1959, *i.e.*, $\frac{225}{145}$. This gives an increase of 55 per cent. Alternatively the product 145 may be termed base 100, then the product 225 is equal to an index of 155, or 55 points higher.

Each of the three 1959 figures calculated so far shows a different increase over 1949. Since the arithmetic is correct, we can only ask which is the 'correct' figure. The simple so-called aggregative index giving an increase of 100 per cent is quite unsatisfactory. If the prices for the various constituent items vary widely in absolute terms the 'index' will always be distorted by

the extreme values. In practice, this method is ignored. The second index required the calculation of the percentage change for each price; by this means the distorting effect of the absolute differences in the prices between the two dates is eliminated. The arithmetic average of the percentage increases might be used for a very simple rough and ready index, but it assumes that each price change is just as important as any other and this, generally speaking, is not true. For example, in a cost of living index a 10 per cent. increase in the price of bread would be much more serious for the family budget than a similar increase in the price of biscuits, even if in absolute money terms the latter increase was greater. But if the percentage changes are 'weighted' in accordance with the relative expenditure on each commodity to the total expenditure, then we get the same results as if we had multiplied the prices at the two dates by the actual quantities purchased. In other words, the *aggregative* index in Table 55, which gave an increase of 55 points, *i.e.*, 1949 = 100 and 1959 = 155, may be directly compared with an index based on the products of the percentage change in each price multiplied by its appropriate share in the total outlay. In this case we are using 'value' instead of 'quantity' as the weight and the 'value' is the amount spent on each commodity in 1949. Table 55 shows the total outlay on each commodity in 1949 and these figures appear in the column headed 'value' weights of Table 56 below.

TABLE 56

METHODS OF CALCULATING INDEX OF WEIGHTED RELATIVES

				Method A		Method B			
Commodity	Prices		'Value' weights	Percentage change in prices	Products of % change & weights	Price Relatives		Relatives × Weights	
	1949	1959				1949	1959	1949	1959
	s.	s. d.							
A	5	7 6	20	+ 50	1000	100	150	2000	3000
B	10	12 6	60	+ 25	1500	100	125	6000	7500
C	15	20 0	45	+ 33⅓	1500	100	133⅓	4500	6000
D	20	60 0	20	+ 200	4000	100	300	2000	6000
	50	100 0	145		8000	400	708⅓	14500	22500

Method A. Weighted percentage change in prices between 1949 and 1959 $\frac{8000}{145}$ = 55 (to nearest unit).

If 1949 prices = 100, 1959 prices = 155.

Method B. Weighted average of 1949 price relatives = 14500

" " " 1959 " " = 22500.

But 1949 prices = 100, then 1959 prices = $\frac{225}{145} \times 100 = 55$ (to nearest unit)

It will be seen that in columns headed 'A', by multiplying the percentage change in each price by the 'value' weight, their product of 8,000 divided by the sum of the value weights 145, gives an increase in the two sets of prices in the period 1949-59 of 55 per cent. The more conventional method is to convert the prices at each date into 'relatives'. Thus if the base year price, *i.e.*, 5s. in 1949, is termed 100, then the corresponding figure for 7s. 6d. in 1959, is 150 or 50 per cent. higher. This figure is referred to as a *price relative*. In the columns headed 'B', each price relative is multiplied by its corresponding weight, and the total products for 1959 expressed as a ratio of 1949. As is expected, the same result is achieved.

Since we have seen that the same result can be derived by either of two methods:

(*a*) an aggregative index using actual prices and quantities purchased as shown in Table 55; or

(*b*) an index of weighted relatives using price relatives and 'value' weights as in Table 56;

the question poses itself, 'which is the better method?' Clearly, the answer is that they are both the same in so far as they produce the same result. From the purely practical point of view the aggregative index seems simpler since it avoids the need for converting prices into relatives. On the other hand, the data available for the construction of the index may be in the form of relatives. This is not so often the case with price indices as with indices such as that measuring industrial production. In this case changes in the output of individual industries may be based on physical output, numbers employed, or on total sales. To combine the changes for a number of industries using different bases of measurement therefore necessitates the use of relatives.

Index Number Notation

The arithmetic processes described so far are usually expressed in simple symbols. Since they are extensively and commonly employed to indicate the type of index used, it is important that the student reader should understand them. If p represents a price, the base year of 1949 is written $_o$ and the year 1959 as $_1$,

then p_0 represents a base year price and p_1 the price of the same article in the other year. The symbol Σ, it will be recalled, is known as large sigma and indicates summation. Thus instead of writing $\frac{p_1^1 + p_1^2 + p_1^3 + p_1^4 + \ldots\ldots p_1^n}{p_0^1 + p_0^2 + p_0^3 + p_0^4 + \ldots\ldots p_0^n}$ where the numbers 1, 2, 3, 4 . . . n represent different goods, we can simply abbreviate to $\frac{\Sigma p_1}{\Sigma p_0}$. This is the formula for the so-called simple aggregative 'index'. But this type is only useful if the prices are weighted by quantities. In our example the quantities were those purchased in the base year, so that if q_0 represents the base year weight, then the index I is merely the average of a series of such weighted prices, *i.e.*, $I = \frac{(p_1 q_0) + (p_1 q_0) + (p_1 q_0) \ldots}{(p_0 q_0) + (p_0 q_0) + (p_0 q_0) \ldots}$ or $\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}$.

The same symbols are used to indicate an index based upon the price relatives. The price relative for any article or goods is derived by relating the base year price to the other year, *i.e.*, $\frac{p_1}{p_0}$. Thus when there are a large number of relatives, we get

$$\frac{p_1^1}{p_0^1} + \frac{p_1^2}{p_0^2} + \frac{p_1^3}{p_0^3} + \frac{p_1^4}{p_0^4} \ldots\ldots \frac{p_1^n}{p_0^n}$$

which can be abbreviated to $\frac{1}{N}\Sigma\frac{p_1}{p_0}$

i.e., the sum of the relatives divided by their number. When the price relatives are weighted, we used 'value' weights. These are given (see Table 56) by the product of the base price (p_0) and the quantity purchased in the base period (q_0). Thus a base year weighted price relative is written $\frac{p_1}{p_0}(p_0 q_0)$ and since the index comprises a number of such weighted relatives, we summarise them by writing $I = \frac{\Sigma\left(\frac{p_1}{p_0} p_0 q_0\right)}{\Sigma p_0 q_0}$.[1]

It was shown in Table 55 that the index derived from the ratio of the aggregate of the *quantity* weighted prices in one year to the corresponding value in the base year gave the same result as the

[1] The student uncertain of this passage should convert the formula back into the actual values used in Table 56.

index obtained by using the *value* weighted price relatives. It must follow therefore that $I = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} = \frac{\Sigma\left(\frac{p_1}{p_0} p_0 q_0\right)}{\Sigma p_0 q_0}$.

Arithmetic or Geometric Mean?

It was stated earlier that an index is merely a form of 'average'. In Chapter VI two types of average were described, the *arithmetic* mean and the *geometric* mean. So far in the foregoing simple illustrations the arithmetic mean has been used, but it is equally possible to use the geometric mean. This statistic is the *n*th root of the product of *n* values. For example, if three numbers are multiplied together, their geometric mean is the cube root of their product; for ten numbers it is the 10th root of their product. We can apply this formula to the data in the above illustration. For purposes of this calculation logarithms are employed, but not merely because they greatly simplify the calculation. Note too that the difference between the logarithms for any two prices of a single commodity gives the anti-logarithm of the relative change. In other words using logarithms saves the trouble of working out the price relatives for every commodity in the index.

Commodity	Prices 1949	Prices 1959	Logarithms 1949	Logarithms 1959	Value weights	Logarithms × Weights 1949	Logarithms × Weights 1959
	s.	s.					
A	5	7·5	0·6990	0·8751	2	1·3980	1·7502
B	10	12·5	1·0000	1·0969	6	6·0000	6·5814
C	15	20	1·1761	1·3010	4·5	5·2925	5·8545
D	20	60	1·3010	1·7782	2	2·6020	3·5564
					14·5	15·2925	17·7425
							15·2925
						14·5)	2·4500
							0·1686

antilog. 0·1686 = 1·47

Since the difference between the logarithms of each price at the two dates gives the relative increase, the appropriate weights are the *value* based weights, not the *quantity* weights. To simplify the arithmetic the value weights in column 6 of the illustration have

simply been divided by ten. The reader will appreciate that it is not the absolute size of the weights which is important, but their ratio one to the other; 6:2 is the same as 60:20. As may be seen in the example the log of each price is multiplied by the corresponding weight. The difference in the sums of the logarithms of these products for each year is then divided by the sum of the *weights*, not the number of prices and this anti-log. gives the relationship between 1959 and the base year 1949. In this case it is 1·47, so that if we call 1949 the base year = 100, then for 1959 the index is 147.

It will be noted that this index is smaller than the index of price relatives based upon the arithmetic mean; this is a characteristic of the geometric mean since it is less affected by the larger values then the arithmetic mean. The obvious question is, which of the two is to be preferred? Both are used. In theory it would be possible to use any 'average', *e.g.*, either the median or mode of a distribution of price changes, as an index. In practice, however, the choice lies between the geometric and arithmetic means of the weighted prices or their relatives. Neither has any marked advantage over the other. Generally speaking, an index based on price relatives often uses the geometric mean while aggregative indices often employ the arithmetic mean, but this is by no means an invariable procedure.

An exposition of the relative merits of these two types of mean lies beyond the scope of this book, for quite apart from the fact that it introduces the highly complex subject of index number theory, it is also subsidiary to the more fundamental problems of index number construction which are now discussed. For the student with some algebra, interested in this aspect of economic statistics, there are several suitable texts for study.[1]

The basic problems of index number construction are three in number. The first concerns the nature of the index, *i.e.*, the functions it is to perform, and the choice of suitable prices or values to be included in it. The second problem is to determine the best weighting system, while the third involves the selection of the base period. None of these particular points is simple to solve in practice and the final index is usually the product of compromise between theoretical standards and the standard attainable with the given data. To simplify the text, references

[1] See in particular appropriate chapters in *Applied General Statistics* by Croxton and Cowden, or *Applied Statistics for Economists* by Karmel, P.H., or *Economic Arithmetic* by Marris, R.

will be to price indices and the base period will usually be assumed to be a year. All the points made are equally relevant to indices measuring other than price changes and where the base period is shorter or longer than a year.

The Nature of the Index

It is helpful to remember when dealing with index numbers that they are specialised tools and as such are most efficient and useful when properly used. A screwdriver is a poor substitute for a chisel, although it may be used as such. All index numbers are designed to measure particular groups of related changes. For example, the Index of Retail Prices in the United Kingdom measures the monthly change in the cost of a collection of goods and services bought by the 'average' household. Note that an index does not cover all such changes, merely a selection. Thus, if one household does not buy fruit at all, when the index goes up because fruit prices rise, that family's real income cannot be determined by reference to that index. The prices of the articles included in the index are those charged at certain types of shop, not every shop. In some cases the allowance for house rent in the index is far below what some families in 'middle class' circumstances pay; just as is the proportion of expenditure on alcohol and tobacco usually well above that applicable to families with young children. In other words, the index is an 'average' of certain household expenditure which determines the standard of life of the average household in the United Kingdom; it does not apply to any single household. Not all households are included in the index; those where the head of the household earns over £20 a week and homes where the old age pension accounts for more than three-quarters of the total income are excluded. Despite all these limitations which are discussed elsewhere (p.270) the public continues to refer to the index as a 'cost of living' index which is indiscriminately applied to all wage and salary earners, whatever their domestic circumstances.

The Board of Trade's Index of Wholesale Prices illustrates the change in attitude to index numbers in recent years. The old pre-war index was designed to measure changes in the 'general level of prices'; it is now recognised that the 'general level' is an over-simplification – it is itself an average of a large number of changes in differing directions and of varying degree. The Board

of Trade therefore now prepares a whole series of indices, each especially constructed to measure small groups of related prices, *e.g.*, the cost of building materials (see Chapter XVII).

Once the decision has been taken as to the purpose of the index the choice of 'representative prices' to be included in the index must be made. The more prices that are included, the longer it takes to calculate the index each month. Furthermore, not all prices are readily available and items may therefore be omitted on this account. Nor is it true to argue that an index with a large number of items is automatically a better index than one based on a few prices. For example, the Board of Trade's old wholesale price index consisted of 258 quotations for 200 commodities; the sensitive index of commodity prices prepared by *The Economist* used only 10 prices. Each served its declared purpose quite well. In the index of Retail Prices only those articles which are regularly and extensively purchased are included. Even so this involves a very large number of price quotations taken from all regions of the United Kingdom, from all types of town and district, as well as from all types of shop![1] How much better for its purpose the index is in consequence of this great amount of detailed work, than if it were to be based on a handful of articles and services from a few shops might well be debated at length! Unfortunately, the government cannot use what might appear to be a 'makeshift' index for this purpose; but it does not follow that the latter would not be so much less satisfactory in the longer period than the present one!

The Choice of Weights

As we have already seen, an index based on a simple average of price relatives or aggregate prices is useless. The individual prices or their relatives must be weighted. What determines the weights? They can be of two kinds. The first are the so-called *quantity* weights used in the aggregative type of index (Table 55) and the others are *value* weights derived by reference to the actual outlay on the particular item which are applied when using the price relatives (Table 56). The actual size of the weights is unimportant; what is important is the *relative* weights. For example, in the current index of Retail Prices the food prices are

[1] As was explained in Chapter XIV.

given a weight of 350 against a weight of 55 for fuel and light. This indicates that food absorbs roughly seven times as much of the average household's weekly expenditure as does fuel and light. Therefore a change in the price of the former is seven times as significant as the same *price* change in the latter. To show the effect on the household's income by means of an index, the food item must be weighted seven times as heavily as that for fuel and light, since the expenditure on the former is seven times as great and takes proportionately more of his weekly income. Whether the weights are 7:1, 350:50 or any other numbers is irrelevant provided they give the true relationship between the items. For example, up to 1947 the old cost of living index weights totalled 100; in the new index they add up to 1,000. That is *not* significant. But the fact that in the old index food accounted for 60 per cent. of the weighting and in the new current index only 35 per cent., *is* important.

The next aspect of the weighting problem is to decide between what is termed base and current year, weighting. For example, in the index of industrial production the weights have been determined by reference to Census of Production data relating to 1954. These weights are applied to output relatives which indicate the changes in monthly output for a wide range of industrial products. It would be possible, but hardly practicable, to weight the relatives by reference to current census of production data derived each year from either a census or sample survey. If current year weighting is used, *e.g.*, 1958, then all the indices for the earlier years must be revised so that they remain comparable one with another and with the latest year. The nature of the weighting is indicated in the formula. For a price index using *base* year weighting we can write $\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}$; the same formula for a *current* year weighted index is written $\frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}$. The reader will observe that the difference is indicated by the subscript q_1, instead of q_0.

The two different weighting systems will give different results, although the actual differences in most cases will not be great. The formula for the base year weighted index is often termed a *Laspeyre* index, and the current year weighted index a *Paasche*

index. These two names have been used for about half a century and are now used to convey to the reader the nature of the index being used. Once again, however, the reader needs to be reminded that the difference between these two types of index is more apparent than real, at least for most practical purposes. In practice, the base year weighting has the great advantage that the weights are constant throughout the life of the index. Furthermore, the work in calculating the index is much less than with current weights. The major disadvantage of base year weighting is that it may become out of date, in which case the efficiency of the index declines. It is for this reason that Laspeyre type indices are revised at fairly regular intervals, although this means that comparability of the index over the longer period is virtually impossible.

The difficulty over base or current year weighting arises because the relative importance of the items comprising the index is continuously changing. If there were no change in the amounts of different commodities bought from year to year, then of course the one original set of weights would serve as both current and base year! Naturally, when prices of the constituent articles in the index change, so does demand and, strictly speaking, the weight needs to be changed. Because people tend to spend less on goods when their prices are rising, the use of the Paasche or current weighting produces an index which tends to understate the rise in prices, just as the Laspeyre index overstates it since the base year weights reflect an outmoded purchasing pattern. But as already stated, the difference between the two formulæ is of interest primarily to the theorist and much has been written on the problem of designing the 'ideal' index. For example, the suggestion has been made that the geometric mean of the Laspeyre and Paasche indices should be used as an index.

In practice, however, the base year weighted Laspeyre type index remains the most popular for reasons of its practicability. The Paasche type index can only be constructed when up to date data for the weights are available. This is exceptional; only one Paasche type index was prepared officially in the United Kingdom – the *average value* index of imports and exports – but even this is no longer published. This was practicable because the monthly overseas trade returns are available with a delay of only a month or two so the weighting could be continually revised.

Selecting the base year

At first sight this should be the easiest part of constructing an index number, not least since it would appear to be logical to take as the base year the first year for which the index is constructed. With a new index this is valid, but with every Laspeyre type index the time comes when the base year must be revised. If the revisions are fairly regular, as for example is the case with the price and volume indices of imports and exports of the United Kingdom, then to some extent the base tends to choose itself. But if the intention is to try and keep a single series of indices for a longer period than a few years, the base period may be difficult to select. This is because the ideal base would be the so-called 'normal' year, but with economic data who is to decide what is normal? At best depression or boom years – in respect of the particular phenomenon measured by the index – can be avoided. Much of the wrangle over the relative success of the two political parties' post-war economic policies could be explained by basing their data on different base years. If a depression year is taken as the base, most later years show an improvement; if a boom year is taken the rate of expansion in the period following appears very slow.

The tendency nowadays is to keep the Laspeyre type of index up to date by regular revision. For example; the index of industrial production was first produced in 1948 with 1946 as a base. In 1952, with the results of the 1948 full census of production available, the index was revised and in 1958 rebased on to 1954 in the light of the census of production data for that year. The index of retail sales has also been revised three times since the end of the war, as well as the import and export trade indices. The reason for the periodic revisions of many indices of economic data is probably the desire to use a reliable index for relatively short period analysis of economic trends. After all, it is only the historian who wants to look back over the longer period; the economist is usually concerned with the immediate short run future. For much historical economic analysis it is usually practicable to link, if only approximately, the old and new index. This can be done quite effectively by calculating for the last two or three years of the old series of index numbers, an index on the new basis. This has been done in the case of the 1958 revision of the index of industrial production, annual

indices have been worked back for a few years to overlap with the old series.

The only index of prices which covers any long period of years is the Sauerbeck index of wholesale prices which is still compiled by *The Statist* in its original form. This dates back over a hundred years, although its base period is 1867-77. This is a rather simple type of index. It covers only basic commodities which do not change in character over the years. But to calculate, as could be done from the records, an index of (say) exports from the United Kingdom for the past century would be pointless. Many goods exported today were unknown in the 19th century, just as many exports of the last century have ceased to be important today. In other words, what purpose would be served by preparing such a long-period index?

One method of overcoming weaknesses in an index from an outdated or frequently changing base is to use a *chain-base* index. In this case, the index for each period is based on the index of the comparable period immediately preceding it. A particular advantage of this type of index is that it is easy to introduce new items. With a fixed base index of the Laspeyre type, to alter the composition of the index would necessitate the re-calculation of the index for all previous years. Against the chain-base type of index is the point that it is really only suitable for the short period. If changes in the component items are frequent, the index may in the later years reflect quite different price movements than the figures in the earlier period.

Conclusions

It will be apparent that an index number, whether it be of prices, of physical quantities or any other measure, is an arbitrary and imperfect measure. At best it will perform the task for which it is designed if every care has been taken to include the relevant constituent items and to weight them correctly. Important though weighting may be, it is still more important to ensure that relevant values are included in the index and that the quotation, *i.e.*, price, from month to month, is comparable, than to devote undue attention to calculating precise weights. Provided the latter are approximately correct, the index should reflect the pattern and trend of change in the data.

Further illustrations

Most index numbers involve many values and much calculation. The highly simplified illustrations on pages 326 and 327 illustrate the basic principles involved, but the actual construction of an index involves a great deal of mechanical work which can hardly be reproduced here. However, the two illustrations which follow may enable the reader to see how two important index numbers are calculated after the basic data have been processed. The first is the famous Sauerbeck index prepared by *The Statist* and the other is the Index of Retail Prices.

The Statist Index of Wholesale Prices comprises 45 commodities divided into two main groups, Food and Materials, each of which in turn consists of three sub-groups of prices. The actual number of prices used to construct each index is given in the first column headed 'Number of Commodities in Index'. There is no weighting by reference to values or quantities but the 'General Average' index is weighted in so far as there are more quotations for certain commodities than for others. The index

TABLE 57

CONSTRUCTION OF 'THE STATIST' INDEX NUMBER FOR 1957*

1866 - 77 = 100

Commodities	Number of Commodities in Index	Total Numbers		Index for 1957
		1867-77	1957	
General average	45	4,500	16,920	376
Food	19	1,900	5,951	313
Vegetable food	8	800	2,356	295
Animal food	7	700	2,402	343
Sugar, coffee and tea ..	4	400	1,193	298
Materials	26	2,600	10,969	422
Minerals	7	700	3,991	570
Textiles	8	800	3,260	408
Sundry materials ..	11	1,100	3,718	338

* *Source: J.R.S.S.*, 1958, Part III, p.348.

for any single commodity is very simple to obtain. The average price for the month is expressed as a price relative of the average price of the commodity in the eleven year base period. Thus, in the index for vegetable food, the price relative for each of the eight commodities is calculated and their average gives the

index for the group. This is repeated for each group in the index. The 'general average' is the simple arithmetic mean of all the 45 price relatives. The monthly indices are based upon the end-of-month quotations, but the annual indices are derived from the average of the 52 weekly quotations, so that the annual index for any year does not necessarily coincide with the average of the 12 months' indices for that year.

The figures under the main column headed 'Total Numbers' are quite clear. For 1867-77 it is apparent that the figure shown is merely the product of the number of commodities each expressed as base 100, *e.g.*, the 'general average' index comprises 45 commodities each of which in the base period was expressed as 100. The relatives for the appropriate commodities in each group having been worked out as described above, they are summed and inserted in the second column under that same heading. The index for 1957 is derived by dividing the number of commodities into the 'total numbers' for those commodities, *e.g.*, Food $= 5{,}951 \div 19 = 313$. This method yields the general average index as well; it is not merely the average of the various group indices.

A full account of the method of the construction of this index, together with annual indices for each group back to the beginning of the century, as well as average prices and monthly indices for recent years, is published annually in the *Journal of the Royal Statistical Society*.

The Index of Retail Prices is fundamentally similar in construction to *The Statist* index described briefly above, although it covers a very much wider range of goods and services. The index is compiled monthly, the prices used being those prevailing on the Tuesday nearest the middle of the month. For each article and service included in the index the current price is expressed as a relative of the price at the base date, 10th January, 1956. As will be seen from Table 58 below, there are ten group indices the weighted average of which is the 'all items' index which is customarily used as a measure of changes in the cost of living. Each group index is separately calculated by deriving the percentage change in the price of each article included in that group, *e.g.*, food, and then weighting the change by reference to the share of the total outlay on that group absorbed by the article. The weighted relatives are then averaged (arithmetic

mean) to yield the group index. Each group index in its turn is then weighted in accordance with its relative importance in the entire household budget. The current weights are shown in the first column of Table 58. The products of the weights and group indices are aggregated and then averaged to give the 'all items' index.

TABLE 58

CALCULATION OF INDEX OF RETAIL PRICES FOR 18TH NOVEMBER, 1958†

Commodity Groups	Weights	Group Index	Weights × Indices
Food	350	108·4	379 400
Alcoholic Drinks	71	105·8	75 118
Tobacco	80	107·8	86 240
Housing	87	124·2	108 054
Fuel and Light	55	116·5	64 075
Durable Household Goods	66	99·9	65 934
Clothing and Footwear	106	102·7	108 862
Transport and Vehicles	68	112·9	76 772
Miscellaneous	59	113·5	66 965
Services	58	115·4	66 932
All Items	1,000	109·8	1098 352

† *Source: Ministry of Labour Gazette*, December, 1958.

The index is published monthly in the *Ministry of Labour Gazette* as well as in the *Monthly Digest of Statistics* in the form of first and third columns of Table 58, *i.e.*, the ten group indices and 'all items' index all expressed to one place of decimals. A detailed account of the construction of this index is published by H.M. Stationery Office.[1]

[1] Index of Retail Prices. Method of Construction and Calculation. H.M.S.O. 1959.

CHAPTER XVII

ECONOMIC STATISTICS

The largest body of published statistical data is assembled as a by-product of the government's daily administration of the economic and social life of the nation. Some of the data are derived from the routine administration of government departments, *e.g.*, the overseas trade statistics from the Customs office, unemployment figures from the Ministry of Labour Employment Exchanges. Other data are derived from specific enquiries conducted by government departments and are provided by members of the public, as on the occasion of the population census, or by the business community as with the census of production.

A particularly useful document for the student contemplating a study of official statistics is the pamphlet prepared by the Treasury in 1953 and entitled 'Government Statistical Services'.[1] This document provides a succinct but overall description of the statistical work in government departments. It explains the origins of many series of data, *e.g.* day to day administration and special enquiries either by census or sample. It also contains some useful comments on the problems which arise when data are to be collected. Apart from an explanation of the legal powers under which information is obtained from industry and the public, there is a short account of the organisation of statistical work in government departments and an account of the origins and functions of the Central Statistical Office. It includes two very useful appendices which alone would justify its publication. The first appendix outlines the various statistical data collected by each government department. The second gives a classification of all published statistics under subject headings, *e.g.* agriculture, education etc. and for each subject shows the principal publications containing the statistics together with a note of their frequency of appearance and the department responsible for producing the data.

[1] Published by H.M. Stationery Office for 1s. 3d.

The unfortunate publicity given to 'form-filling in triplicate' has perhaps tended to conceal the positive benefits of these activities. For every branch of economic activity there is a mass of valuable data which should interest the business man and industrialist. Two studies which have been described as of particular interest to every manufacturer and salesman are based almost entirely on official statistics.[1] Yet few of the readers of these books would have thought of studying the various Blue Books and White Papers in which the data first appeared. A particular disadvantage of official statistics, quite apart from the fact that the layman is usually ignorant of their existence, lies in the fact that often they have been assembled for a special and limited purpose and for other uses they are inadequate. A good example of the latter is given by the official labour statistics: until 1948, the government always knew the number of unemployed insured workers, but it could only estimate the size of the working population.

The need to control the economy during the last war and the generally accepted case for a degree of economic planning have brought home to the government in no uncertain manner the deficiencies in their statistical information. For example, without full knowledge of the labour force, control of the distribution of workers was impossible. Similarly, ignorance of the costs of retail distribution make estimates of the national product little better than mere guesses. With the support of the government an inter-departmental committee has been created and entrusted with two tasks.[2] The first is to recommend any improvements that may be made in the existing coverage of *official* statistics. The second is to assist in explaining what data are available on various subjects. Progress in this respect is regrettably slow; only four studies have so far been published.[3] The Royal Statistical Society, however, has produced a series of papers covering a wide range of subjects, from the census of production to the statistics of brewing, which indicate the main sources of data, *both* official and private.[4]

[1] The Home Market, 1950. Revised edition. M. Abrams. Allen & Unwin. Marketing Survey of the United Kingdom. Business Publications Ltd., 1951.

[2] Interdepartmental Committee on Social and Economic Research.

[3] Guides to Official Sources, No. 1, Labour Statistics. No. 2, Census Reports of Great Britain, 1801-1931. No. 3. Local Government Statistics. No. 4, Agriculture and Food Statistics. Published by H.M. Stationery Office.

[4] The Sources and Nature of Statistics of the United Kingdom. Vol. I and Vol. II.

The purpose of this chapter is to indicate particular sources which are of interest to the student of economic affairs. No attempt is made to be comprehensive or to discuss the detail of such statistics. Emphasis will be laid on the scope and any deficiencies of the existing data.

The most useful source of statistical information compiled by the government is the annual 'Abstract of Statistics'. This is a joint publication by the Central Statistical Office and the Statistical Divisions of the various government departments. Although most of the data are collected by government agencies, some information is provided by private organisations. The 'Abstract' is published annually and for many series the data for a ten-year period are brought together. As a general rule, when any information is required, the annual Abstract should be the first source to be consulted. Since 1946 the Central Statistical Office has produced the *Monthly Digest of Statistics*, which, like the Annual Abstract, has a wide coverage. It excludes, however, the data relating to social conditions and deals primarily with economic data. Monthly figures of supplies of fuel, raw materials and finished products together with indices of output in selected manufacturing industries are provided. These are supplemented by data covering labour, wages, transport, and foreign trade, as well as important financial statistics. For most of these items the *Monthly Digest* gives monthly data for the last one or two years, together with comparable figures for the earlier years. To ensure that the data given in the *Digest* are correctly interpreted, a supplement is issued in January each year. This provides detailed definitions of units and items given in the *Monthly Digest*.

Another useful monthly publication entitled *Economic Trends* is produced by the Central Statistical Office in collaboration with the Statistical Divisions of government departments. It provides both charts and statistics illustrating current trends in the United Kingdom economy. Each issue also now contains at least one article commenting on features of current economic statistics or introducing a new series, or describing methods used by a Statistics Division of a government department in the preparation of their particular data.

More detailed information on a very large variety of economic subjects is published at intervals in the weekly *Board of Trade Journal*. For current data this is more up-to-date and

useful than either the *Annual Abstract* or the *Monthly Digest*. Much of the industrial and commercial information contained in the *Monthly Digest* appears in rather greater detail and earlier in the *Board of Trade Journal*. Important in its specialist field is the *Ministry of Labour Gazette* published monthly. All data relating to the labour force, working conditions, and wages, are first published in the *Gazette*.

We now go on to consider a few selected and more important sections of official statistics.

Manpower

In the situation described as 'full employment', the government must pay special attention to statistics of labour. Despite the fact that the size of the national income of the community is ultimately dependent on the size and productivity of its labour force, it is only since mid-1948 that reliable estimates of the working population have been prepared. Until the beginning of the war, the size of the working population could only be estimated, although the number of insured unemployed was known fairly accurately. All estimates of the working population made until 1939 were based on the insured population, which since the inception of the Unemployment Insurance scheme in 1911, had included only part of the working population. These annual totals of 'insured workers' were not, in fact, really comparable, since the scope of the Insurance Acts was changed at intervals to include new groups of workers. Thus the insured population of some 10 million in 1920 had risen to over 18 million by 1939. The percentage figures of unemployment which were calculated monthly were based on the number of insured workers registered as unemployed, *i.e.*, those who had lodged their cards at the local employment exchange and the total number of insured workers.

To calculate the size of the total labour force, those sections of the working community outside the scope of the Acts had to be estimated. The largest section were non-manual employees in receipt of more than £250 per annum,[1] employers and the self-employed. Furthermore, the number of private domestic servants could only be tentatively estimated. In any case, comparability between the monthly totals of insured workers before and

[1] Raised in 1940 to £420 p.a.

after September 1937, is affected by the introduction of a revised counting procedure. Before the war only the decennial Census of Population held in 1931 yielded comprehensive data on the working population.[1] The distribution of workers actually employed in manufacturing and extractive industries was also derived from the 1935 Census of Production. Unfortunately, the classification employed in both of these censuses differed from that of the Ministry of Labour. This situation no longer exists since the introduction of the Standard Industrial Classification in 1948. All government departments now use this as the basis of industrial classification.

The wartime problem of allocating scarce labour between competing claims led the government to fill in the serious gaps that existed in its labour statistics. Quarterly returns of employees from firms in the engineering industry were introduced early in 1940. Half-yearly returns were collected from the textile industry, as well as from a sample of firms in the catering and distributive trades. These figures were supplemented by official returns of staff and employees in the Civil Service and Local Government. Regular returns of certain classes of workers were made by various government departments, *e.g.*, the number of building operatives was returned by the Ministry of Works, and that for teachers by the local authorities. By the end of the war, the Ministry of Labour was in a position to make fairly reliable estimates of the working population. The major source of weakness in the figures arose from the lack of information about the numbers in the employer and self-employed groups. Nevertheless, comparable published estimates of the working population compiled on this basis are available for June 1939, and for the same month in every year from 1941 to 1947. The steady improvement in the coverage of the data collected enabled the Ministry of Labour to publish comparable figures for *every* month from mid-1945 to January 1949, revealing the distribution of the labour force between a very large number of industries.

The main lesson to be learnt from the above sections is that the continuous extension in the coverage of the insurance figures before 1939 makes impossible inter-year comparison of totals for more than a few years at a time. Whatever source is

[1] For example, the 1931 Census Report contains the only comprehensive data relating to the professions. Estimates could, of course, be made from the membership lists of the professional societies.

consulted, a careful check of the footnotes is essential if the data are to be correctly interpreted.

The introduction in July 1948, of the National Insurance Act resulted in the provision of new and comprehensive data. Under this Act, every gainfully occupied individual must register. In theory, therefore, the totals of insured persons should yield an exact figure of the labour force of this country.[1] Unfortunately, there is ample evidence that some employers and self-employed persons are not registered under the Act. There are, too, at the present time a number of aged workers beyond its scope. When the scheme was introduced, every worker received an insurance card, and the count of these cards provided the first reliable estimate of the entire working population, subject to the remarks above. Due partly to the greater coverage of the 1948 Act and to the counting of part-time workers as whole units instead of half-units as was done before, the figure for the 'working population' at mid-1948, as given in the published statistics, rose by about two million.

With an insured population of some 23 million persons, it is not administratively possible to continue, as under the old Act, the practice of an annual exchange of all cards at one time, *i.e.* in July each year. Instead, the cards are now exchanged in four groups at the beginning of March, June, September, and December. Each group, however, constitutes a random sample of the population (of cards), and tentative estimates of the total number of insured workers and changes therein can be obtained by multiplying by four any one of the four quarterly totals of cards exchanged. It is not practicable, however, to use a single quarter's cards for an *analysis* of the industrial distribution of the labour force. This is done once yearly. The estimated distribution at end-May is obtained by supplementing the data derived from the June exchange of cards with returns from employers of 5 or more workpeople in June, indicating the number of cards held by them, and the number due for exchange in that month. This enquiry covers more than three-quarters of the employed population. These returns coupled with the June cards actually returned, are analysed on the basis of the Standard Industrial Classification and the final figures are published

[1] Not all married women in employment are included since they are covered for limited benefits by their husband's insurance.

annually – usually in February following the count in the *Ministry of Labour Gazette*.[1]

The end-May totals of employees calculated as described above, and published in the *Gazette* of the following February, relate to all industries in both Great Britain and the United Kingdom. The totals are analysed by sex and by age; in the case of the latter it is a breakdown as between the under and over 18 year olds for both sexes. These estimates include not only those employees absent from work on account of sickness or any other cause but those in the various industries who are unemployed. The June issue of the *Gazette* contains a regional analysis of the total number of employees classified by age and sex, together with an estimate of inter-regional migration of each age/sex group, as well as a detailed age analysis of the employees in each industry. In addition there is an estimate of the number of married women in employment. In the August issue there is published an important set of data showing the number of young persons under 18 years of age taking up employment for the first time during that year, classified by the industry entered, age, and the type of employment taken up.

TABLE 59

TOTAL WORKING POPULATION – GREAT BRITAIN

	Strength (in thousands) at		Change during twelve months June 1957-58	
	June 1957	June 1958	'000's	Per cent
Total Working Population ..	24,188	24,070	— 118	— 0·5
of which:				
Men	16,225	16,166	— 59	— 0·4
Women	7,963	7,904	— 59	— 0·7
H.M. Forces and Women's Services	702	614	— 88	— 12·5
Total in Civil Employment ..	23,245	23,080	— 165	— 0·7
of which:				
Men	15,367	15,294	— 73	— 0·5
Women	7,878	7,786	— 92	— 1·2
Registered Unemployed* ..	250	432	+ 182	+ 72·8
of which				
Wholly Unemployed* ..	235	370	+ 135	+ 57·4
Temporarily Stopped* ..	15	62	+ 47	+ 313·3

Source: Ministry of Labour Gazette, August 1958.

* End of month estimates. Persons classed as temporarily stopped are included in the totals of persons in civil employment.

[1] For a detailed account of the counting procedure and the final analysis see the *Gazette* for February 1958.

Apart from the foregoing annual figures, the *Gazette* also publishes various quarterly statistics. One set of data measure the labour turnover in the manufacturing industries – intake and wastage being expressed as percentages of the total number of employees. The other provides a summary of the information supplied by employers in manufacturing industry about short-time and over-time working. These figures are based upon a quarterly return from employers and are published for each of the main industrial *groups*, as well as for selected individual industries within each group. The group and industry statistics relate to the number of workpeople actually working either short or over-time, as well as an estimate of the aggregate number of hours lost or overtime worked.

The best known figures published in the *Gazette* are the monthly analyses showing the numbers of men and women

TABLE 60

MANPOWER IN CIVIL EMPLOYMENT IN GREAT BRITAIN

Industry or Service	Strength (in thousands) at		Change during twelve months June 1957–58	
	June 1957	June 1958	'000's	Per cent
Agriculture and Fishing ..	1,025	1,002	— 23	— 2·2
Mining and Quarrying ..	868	854	— 14	— 1·6
Chemicals and Allied Trades..	534	529	— 5	— 0·9
Metal Manufacture	579	558	— 21	— 3·6
Vehicles	1,225	1,241	+ 16	+ 1·3
Engineering, Metal Goods and Precision Instruments ..	2,814	2,785	— 29	— 1·0
Textiles	934	864	— 70	— 7·5
Clothing (inc. footwear) ..	678	648	— 30	— 4·4
Food, Drink and Tobacco ..	916	929	+ 13	+ 1·4
Other Manufacturers ..	1,591	1,565	— 26	— 1·6
Total in Manufacturing Industries	9,271	9,119	— 152	— 1·6
Building and Contracting ..	1,519	1,495	— 24	— 1·6
Gas, Electricity and Water ..	379	378	— 1	— 0·3
Transport and Communication	1,723	1,715	— 8	— 0·5
Distributive Trades	2,945	2,979	+ 34	+ 1·2
Professional Financial and Miscellaneous Services ..	4,217	4,247	+ 30	+ 0·7
National Government Service	543	530	— 13	— 2·4
Local Government Service ..	755	761	+ 6	+ 0·8
Total in Civil Employment ..	23,245	23,080	— 165	— 0·7

Source: Ministry of Labour Gazette, August 1958.

comprising the total working population, *i.e.*, those in civil employment, the unemployed and members of the forces. The form of publication is reproduced in Table 59. This is supplemented by an analysis of manpower in civil employment into the main industrial groups as shown in Table 60. These data are in their turn given in more detail for individual industries where monthly data are available, *i.e.*, from all firms employing 100 or more, and from a 25 per cent sample of those firms with between 11–99 employees. These tables classify the labour force by sex and show the actual numbers employed in each industry at selected dates in each of the quarters.

Apart from data on the employed population, detailed statistics about the unemployed insured population are prepared. Monthly analyses of these data are given, classified by age and sex, by industry, by duration of unemployment, and by regions and principal towns. Since the war, analyses of the unemployed totals in the Development Areas have also been published. Table 61 shows one of the monthly tables on unemployment. Since 1948 all the data have been classified on the basis of the Standard Industrial Classification, so that comparisons of current data, *e.g.*, of wage rates in particular industries, with the

TABLE 61

NUMBERS AND RATES OF REGISTERED UNEMPLOYED – GREAT BRITAIN

Region	Numbers of Persons Registered as unemployed at 13th October 1958			Percentage rate of unemployment (*a*)		
	Males	Females	Total	Males	Females	Total
London and South-Eastern ..	55,905	20,428	76,333	1·6	1·0	1·4
Eastern and Southern	28,526	9,922	38,448	1·9	1·3	1·7
South Western ..	21,706	8,266	29,972	2·7	2·1	2·5
Midland	29,008	11,899	40,907	2·1	1·6	1·9
North-Midland ..	18,779	7,371	26,150	1·9	1·5	1·7
East and West Ridings	32,070	13,876	45,946	2·6	2·2	2·5
North-Western ..	60,297	35,275	95,572	3·2	3·2	3·2
Northern	25,461	10,294	35,755	2·8	2·7	2·8
Scotland	60,815	25,002	85,817	4·3	3·3	4·0
Wales	27,188	11,754	38,942	4·0	4·4	4·1
Great Britain ..	359,755	154,087	513,842	2·5	2·0	2·3

(*a*) Number registered as unemployed expressed as percentage of the estimated total number of employees.

Source: Ministry of Labour Gazette, November 1958.

figures from earlier years are not always possible. This problem has been partly overcome by the publication of specially compiled series back to 1947 of earnings for each class of worker and for selected industries. These 'historical' tables appear half-yearly. See for example the tables in the *Gazette* for September 1958.

All the data and tables published in the *Gazette* are accompanied by explanatory notes which indicate briefly the source and composition of the figures as well as any weaknesses affecting comparability in the series over time. This is not the case when these figures are reproduced elsewhere, for example, in *Economic Trends* or the *Monthly Digest*. If any calculations or comparisons are to be made with published data on labour matters, the reader is well advised to extract his data from the primary source. For the research worker collating data over a long period, a study of the official pamphlet given in the references below is essential.

REFERENCES

Labour Statistics. Guides to Official Sources No. 1. H.M.S.O.

Production

The importance for the national economy of knowing the level of production and its composition needs no emphasis at the present time. Many industries publish data relating to their activities, *e.g.* the post-war Ministry of Power publishes quarterly statements on the coal mining industry, as well as an annual digest of statistics on all forms of fuel production and consumption. *Lloyd's Register of Shipping* provides an annual return of all ships over 100 tons gross under construction in the United Kingdom. The iron and steel industry publishes a monthly bulletin of statistics containing figures relating to the level of employment, output of various products, prices, international trade, and foreign production. Invaluable as these published data undoubtedly are, they relate only to segments of the national economy. The only way to find out the total value of all production in the country is to carry out a census of production.

Apart from this, such a census provides a great deal of information on other points. It reveals the division of the national industrial product between the various industries. The changes in

these data over time bring out the trend and relative importance of the individual industries. Without such information, central economic planning in respect of the distribution of labour and new capital construction is virtually impossible. Estimates can also be made of labour productivity and the ratio of supervisory staff to operatives in the different industries although such figures are of limited accuracy and value. Without all these data, the index of industrial production would be unreliable and estimates of the national product subject to wide margins of error.

The first census of production in the United Kingdom was taken in 1907, and was followed by others in 1912, 1924 and 1930. The last pre-war census, known as the fifth census, was held in 1935. Since the end of the war, following upon the Statistics of Trade Act 1947, a partial census has been taken in respect of industry in 1946, while a full census was held in 1949 relating to industry in 1948. A census was taken for the years 1949 and 1950, but the information then required was rather less than was required in the full census covering the year 1948. The 1948 census was restricted to Great Britain, *i.e.*, no census was taken in Northern Ireland, but with the passage there of an Act similar to the 1947 Act in this country, censuses were taken in Northern Ireland for the years 1949-51 inclusive and the results incorporated with those for Great Britain in the appropriate Board of Trade census reports for those years. It was intended to hold an annual census of production as from 1948 onwards. In the event, a complete census like that of 1948, in which a great deal of detailed information was required from firms, is to be carried out only once every three years. Thus the 1951 census was a full-scale detailed census similar to that of 1948, while a similar census was taken for the year 1954. The Verdon Smith Committee recommended another full-scale census relating to the year 1957 which the President of the Board of Trade deferred and it will now relate to 1958.[1] For the intervening periods, *i.e.* the years 1952 and 1953 and the years between 1954 and 1958, the Board of Trade has conducted either sample surveys of industry or full censuses using a modified schedule with fewer questions. In the sample surveys, returns will be required from about one in seven of all establishments covered

[1] See Schedules reproduced on pp. 27-33.

by the census, and no return at all is made by the 80,000 manufacturing firms employing 10 or fewer persons. The samples cover some 35,000 industrial establishments instead of the quarter of a million or so covered by the census. The sample enquiries differ from the full census in that no detailed questions are asked about goods sold or materials purchased. The questionnaire has been reduced to a single sheet for these sample enquiries.

The data collected in the census of production do not always relate to the calendar year, although we write about the '1948' census. To facilitate the completion of the schedules, the 'establishment' or firm may give figures relating to its *financial* year and not the calendar year. The effect of this concession is that '1957' for example, can mean any twelvemonth period ending between 6th April 1957 and 5th April 1958. According to Mr H. Leak, a former director of the census, the mean year-end of the reporting firms is mid-December.

For statistical purposes the term 'production' requires careful definition. Thus, from the economic point of view, any goods or services produced and exchanged for value constitute 'production'. The census, however, is restricted to the extractive, building and manufacturing industries in both private and public ownership. The first category includes mining and quarrying, but not agriculture, the last group includes firms which are engaged in repair work for the trade, *e.g.*, a ship-repairer. Despite the use of the term 'census', the enumeration of firms is far from complete. In Great Britain only those firms employing more than ten workers return the full schedule. For smaller firms, *i.e.*, those with ten or fewer workers, a return giving the nature of the trade carried on and the number of employees only is required. In certain trades, however, in which small firms are believed to represent a large proportion of total output the 1948, 1951 and 1954 full censuses required such firms to make a simplified return. This varied from trade to trade. Such establishments are known as 'small firms'; it should be noted that the data derived from the Northern Ireland censuses does not include any information relating to small firms. The census there covers only firms employing on the average more than 10 persons, described in the census as 'larger establishments'.

The Board of Trade publishes reports on each industry

TABLE 62

CENSUS OF PRODUCTION: BREWING AND MALTING INDUSTRY

Summary of returns received from firms employing on average more than 10 persons

Item		Unit	Great Britain 1948	United Kingdom 1951	United Kingdom 1954
1	Number of establishments ..	No.	695	623	712
2	Total value of sales and work done ..	£'000	465,603	440,825	450,732
3	Products on hand for sale and work in progress: at beginning of year	,,	17,177	20,043	22,332
4	Products on hand for sale and work in progress: change during year	,,	+ 3,082	+ 1,368	— 277
5	Gross output (production) ..	,,	465,390	442,193	450,456
6	Purchases of materials and fuel ..	,,	81,450	102,036	96,944
7	Stocks of materials and fuel: at beginning of year	,,	17,185	20,979	27,387
8	Stocks of materials and fuel: change during year	,,	+ 2,916	+ 5,066	— 3,166
9	Costs of materials and fuel used ..	,,	78,534	96,970	100,111
10	Payment for work done on materials given out	,,	415	364	429
11	Customs and Excise duties: on beer brewed (net)..	,,	289,979	247,749	238,177
12	Customs and Excise duties: on deliveries for home consumption of wines and spirits	,,	5,645	7,249	5,843
13	Payment for transport	,,	3,294	4,736	5,388
14	Net output	,,	90,816	85,125	100,509
15	Average number of employees: operatives ..	No.	56,621	55,767	52,712
16	Average number of employees: others	,,	14,413	15,058	15,063
17	Total employment	,,	71,069	70,847	67,792
18	Net output per person employed ..	£	1,278	1,202	1,483
19	Wages and salaries: of operatives	£'000	16,132	18,274	21,008
20	Wages and salaries: others	,,	7,914	9,016	10,182
	Capital Expenditure:				
21	New building work	,,	996	1,546	2,266
22	Plant and machinery: acquisitions	,,	4,206	6,048	4,296
23	Plant and machinery: disposals	,,	118	131	381
24	Vehicles: acquisitions	,,	1,139	1,110	1,175
25	Vehicles: disposals	,,	105	125	228

Summary of returns received from firms employing on average 10 or fewer persons

	Unit	Great Britain 1948	United Kingdom 1951	United Kingdom 1954
Number of returns	No.	112	100	92
Total employment including working proprietors	,,	568	541	516

Source: Census of Production 1954. Vol 9: H.

Additions may not agree due to rounding of figures.

covered in the census. In addition, for each census at the commencement of the publication of these industry reports, a document entitled 'Introductory Notes' is produced. This contains a detailed account of the scope and scale of the census together with the definitions employed as well as an explanation of the tables published in the individual industry reports. Whenever data are to be extracted from the industry reports, although the latter contain notes relating to the tables, it is advisable to turn to the fuller 'Introductory Notes' to avoid errors in extraction.

Table 62 is taken from the industry report on those establishments engaged wholly or mainly in brewing and malting. The table it will be noted deals mainly with the larger establishments, *i.e.* those with more than ten employees; information is given, however, of the number of smaller firms and their total employees. The terms used in Table 62 are carefully defined in each report as follows:

(1) The *gross output* of a trade is the total value of the goods made and other work done during the year. It is in effect the value of sales together with the change over the year in the stocks of work in progress and finished goods. This is shown in the table as items 2 plus 4 equals 5, or gross output. It will be seen that this relationship of items does not stand in the case of 1948. This is because before 1951 the figure of gross output excluded any payment for carriage inwards charged on raw materials.

(2) *Cost of materials and fuel used* are inflated after 1948 by the inclusion of transport costs, which were excluded in the previous years. The term 'materials' covers all manufacturing costs from consumable tools and plant repairs to packing materials. In a large organisation the final product of a separate manufacturing unit is often the raw material of another unit. These independent 'units' usually make separate returns to the census authorities, but there has been a change in the method of valuing materials transferred in this way. Since the war the value is to be that which would be charged to an outside purchaser. In the pre-war censuses some firms charged out the goods at their internal costing price, *i.e.*, a fictitious value employed for internal finance and costing. It will be seen from Table 62 that the cost of materials and fuel *actually used* during the year is the figure (item 9) calculated from the purchases and stocks of fuel which are shown separately (items 6–8).

(3) The *net output* constitutes the firm's contribution to the national product, *i.e.* the value added to the materials by the manufacturing process carried out by this firm. Alternatively, net output may be regarded as the fund from which wages, salaries, rent, depreciation and all selling expenses as well as profits are made. As well as deducting the cost of materials and fuel used (item 9) from the gross output, payment for

sub-contracted work (item 10) and transport (item 13) as well as Customs and Excise duties (items 11-12) are also taken away. The resulting amount is termed the net output (item 14).

(4) The table contains information relating to wages and salaries of persons employed as well as the number of employees. On the basis of these two items a figure described as *net output per person employed* is derived. This particular figure must be interpreted with great care. It is at best a poor indicator of the relative efficiency of labour in different industries. An obvious and important factor affecting output as between different industries is the degree of mechanisation in them. There is, too, the often overlooked fact that the final products, especially when the comparisons are between different countries, are seldom identical. Contrast, for example, the British and American 'family' car. This particular figure serves primarily to indicate the changing productivity of labour within the industry. If we assume that the working week remains unchanged then any increased productivity per worker is presumably accounted for either by a greater degree of mechanisation or a more intensive and efficient utilisation of resources.

(5) It will be noted that figures are given for three years. These three sets of figures, however, are not strictly comparable as the 1948 census covered Great Britain only, not Northern Ireland. The 1951 and 1954 censuses, however, covered the United Kingdom, *i.e.* they included Northern Ireland. Care must be exercised, as always, when extracting such data. Whenever in the official Board of Trade publications data for these three censuses are given without reference to their different coverage, then it may be assumed that any adjustments necessary as between the results of different years have been made. For the person unacquainted with the many changes in the Censuses of Production, the danger arises that in an attempt to show changes over a longer period, for example as between pre-war and post-war years, for any industry, figures which are not in fact comparable may be extracted and set side by side as if they were.

Table 63 is another standard table reproduced in all the individual industry reports. It provides an analysis by size, *i.e.*

labour force, of the larger establishments. The student will note that the class interval in the first column headed 'average number employed' is not constant. It is clearly unnecessary that it should be so, since the information provided in the Table is sufficiently detailed for all practical purposes. The information relating to remuneration of operatives and other employees is a feature of the post-war censuses. Pre-war, the size of the firms' labour force alone had to be returned. The classification then employed consisted of two groups: operatives, covering all manual workers; and administrative, technical and clerical staff. Both of these two groups were further sub-divided as either 'over' or 'below' 18 years of age. It was only with the passage of the Statistics of Trade Act 1947 that firms were compelled to make a return of salaries and wages. This information was first available in the 1948 census.

As mentioned earlier, sample censuses of production are carried out in the years between the full censuses, *e.g.* 1955, 1956 and 1957. While not achieving the absolute accuracy of a full census, the sample covers a large proportion of the industry. For instance, the 1957 sample although drawn from about one in nine industrial establishments, included all of the largest ones and accounted for about 70 per cent of the total output. In certain industries, *e.g.* coal mining, gas and electricity supply, etc., sampling methods were not necessary as full information was already available. Sampling was also unnecessary in Northern Ireland. A sampling frame stratified for each industry is employed and the provisional figures which are first published in the *Board of Trade Journal* are unlikely to be substantially amended. An example of the information given in the *Board of Trade Journal* is shown in Table 64. Similar information is given for nearly all the industries covered by the full census. Note that the three years prior to 1957 are given for comparative purposes. These sample figures refer to the entire industry, not just the larger establishments.

A particular feature of the Census of Production data is that the classification employed is the Standard Industrial Classification, which is also employed by the Ministry of Labour so that data from both sources relating to labour distribution between industries and different occupations are now comparable. This was a considerable advance upon the situation

TABLE 63

1954 CENSUS OF PRODUCTION: BREWING AND MALTING INDUSTRY

Larger establishments analysed by size of labour employed

Average number employed	Establish-ments	Gross Output	Net Output	Employees		Wages and Salaries		Capital Expenditure	Net output per persons employed
				Operatives	Others	Operatives	Others		
	Number	£'000	£'000	Number	Number	£'000	£'000	£'000	£
11 - 24	294	17,199	3,971	2,963	628	1,181	437	348	1,103
25 - 49	123	25,624	5,852	3,654	906	1,427	614	438	1,281
50 - 99	130	60,832	12,856	7,227	2,046	2,769	1,315	1,291	1,386
100 - 199	79	66,321	17,014	8,647	2,673	3,227	1,685	1,203	1,503
200 - 299	34	55,296	11,929	6,663	1,900	2,529	1,204	721	1,393
300 - 399	18	46,120	10,294	4,697	1,323	1,797	847	775	1,710
400 - 499	14	48,165	9,801	4,748	1,422	1,974	880	1,261	1,589
500 - 749	7	28,468	6,935	3,334	1,016	1,279	711	215	1,594
750 - 999	6	33,521	7,586	4,150	1,074	1,770	837	458	1,452
1,000 - 1,999	7	68,909	14,270	6,629	2,075	3,055	1,653	1,057	1,639
Total ..	712	450,456	100,509	52,712	15,063	21,008	10,182	7,737	1,483

Source: Census of Production 1954, Vol. 9, H.

TABLE 64

CENSUS OF PRODUCTION 1957: BREWING AND MALTING INDUSTRY

	Unit	1954	1955	1956	1957
Gross Output	£m	454·1	471·9	494·7	520·3
Net Output	„	101·3	107·7	113·7	127·1
Average number employed:					
Total including working proprietors	Thousands	68·3	69·7	67·6	68·9
Operatives..	„	53·1	54·0	52·0	52·8
Other employees	„	15·2	15·7	15·6	16·1
Wages and salaries:					
Operatives..	£m	21·2	22·9	24·1	25·9
Other employees	„	10·3	10·9	11·4	12·5
Change during the year in:					
Stocks of materials and fuels..	„	— 3·2	— 0·2	+ 0·5	+ 0·5
Work in progress and stocks of products	„	— 0·3	— 0·5	+ 0·7	+ 1·1
Capital expenditure:					
Plant, machinery and vehicles acquired	„	5·6	5·8	6·5	8·7
New building work	„	2·3	3·0	3·2	3·7

Source: Board of Trade Journal, 21st November 1958.

which existed before the Standard Industrial Classification was introduced in 1948. Prior to that date, data relating to the labour force, in all cases incomplete, were obtainable from the Census of Production, the Census of Population, and the Ministry of Labour Statistics. The data from each source were classified on different bases.

The Census of Production is the most valuable source of information available to the Central Statistical Office for computing the national product. It is particularly important in view of the information it gives relating to changes in stocks of both finished products and work in progress, as well as of materials and fuel. The information provided by the recent Census of Production relating to capital expenditure on plant, machinery and vehicles, as well as new building work, is a big step forward in filling a very serious gap in our national economic statistics.[1]

The census, however, did not fully meet the requirements and a quarterly sample of over 600 companies covering the manufacturing, distributive and service and transport industries was instituted at the beginning of 1956 to give details of capital expenditure. The results of these sample enquiries are published quarterly in the *Board of Trade Journal*. These figures not only

[1] See pp. 27-33 for the type of schedules used in the Census; they will indicate the scope of the information collected.

show the amount of investment outlays in past periods, but also give an indication of investment intentions for the coming period. These 'forward looking' statistics, as they are termed, are very important. They are limited as yet, having only recently been started in order to simplify the task of forecasting economic changes. Too much reliance should not be placed on the absolute figures. Rather it is the underlying movement from quarter to quarter and year to year which is significant; *e.g.* a sharp drop after several periods of rising outlays may presage a loss of business confidence which, if not countered by government spending, will lead to a deflationary spiral.

In the sample Census of Distribution which was undertaken in respect of 1957, questions on capital expenditure were asked. The Board of Trade is already aware of much of the building and civil engineering work carried out for public authorities and nationalised industries. There was, however, a serious gap in respect of private housing, of smaller works, and of the rate at which contracts were met. To cover this gap the Ministry of Works is collecting information quarterly from the building industry.

A similar position exists in respect of trading stocks held by manufacturers and traders. The Census of Production provides much useful information but it needs supplementing. The Board of Trade began a quarterly enquiry in 1953 which now covers about 350 manufacturers who account for over 40 per cent of the total value of stocks held by manufacturing industry. Even this, however, did not cover stocks held by distributive trades and the Board of Trade have started two enquiries to extend their information in this field. The first is a quarterly sample of wholesalers' stocks and the second is an annual sample of stocks and fixed investment in the distributive and service trades. The first enquiry related to the situation in 1956.[1]

It will be seen that much has been done since 1945 to provide comprehensive and up-to-date information on those sectors of the national economy which are covered by the censuses of production and distribution.

REFERENCES

Report of the Committee on the Censuses of Production and Distribution, Cmd. 9276, H.M.S.O. 1954.

[1] A description of the methods of estimating both annual and quarterly figures of stocks is given in Economic Trends, March 1959.

Censuses of Production and Distribution, H. Leak in Royal Statistical Society's Volume I 'Sources and Nature of the Statistics of the U.K.'

The Report on the Census of Production for 1954. Introductory Notes. H.M.S.O.

Index of Industrial Production

The purpose of the Index of Industrial Production is to provide a general measure of monthly changes in the level of industrial production in the United Kingdom. The index is prepared by the Central Statistical Office in collaboration with the various statistical divisions of certain Ministries, in particular the Board of Trade and the Ministries of Supply and Works. The index is published monthly in the *Monthly Digest of Statistics* and the *Board of Trade Journal*. An official account of the construction of the index following the 1952 revision is available.[1]

The index of industrial production covers mining and quarrying, manufacturing, building, and public utilities, gas, electricity and water, but excludes agriculture and transport. While the index is designed to reflect changes in the level of industrial production from month to month, the individual series or indicators are based as far as possible on weekly rates of production in each industry. Some of the industries and industrial groups for which separate indices are prepared are shown in Table 67. It covers production in both private firms and nationalised industries, as well as central and local government establishments. The index incorporates about 1,300 individual series, each series being separately weighted. By 'series' is meant the indicator relating to output in individual industries. As may be seen from Table 65 over four-fifths of the indicators used to measure production from industries covered by the index are based on output data. But this overstates the extent to which the index is based on figures of output. Owing to inadequate data for certain series, it has been necessary to use as indicators of production data relating to the value of deliveries or sales; in some cases reliance is placed upon figures of the material consumed and, occasionally, the labour employed or man-hours worked. These alternatives are not as satisfactory as the output data itself, but basically it is true to say that the index is based largely on series of physical output. The following table, taken

[1] The Index of Industrial Production. Studies in Official Statistics No.2. H.M.S.O.

from the official account of the index, illustrates the sources of production indicators used in compiling the index together with the percentage weighting attached to each category of indicator. As a result of the 1958 revision when the index was recalculated on to a new base, 1954, for which year new census of production data were available, some additional and improved series or

TABLE 65

Nature of Series used in Index revised in 1952. Base Year 1948.	No. of Series	Percentage of weight carried
Output		
Quantities delivered or produced	1,150	55·2
Value of deliveries or sales	100	26·5
Input		
Quantities of major materials received	37	11·7
Number of persons employed	13	6·6
	1,300	100·0

Source: H.M.S.O. report. op. cit.

'indicators' were introduced. For example, indicators based on 'values of deliveries or sales' which provided 26·5 per cent are now available for 32 per cent of the total weighting, while the 'number of persons employed' indicators now account for only 2½ per cent of the weighting. A further improvement in the index is possible since the indicators expressed in value terms can now be satisfactorily 'deflated', *i.e.* adjusted to reflect volume changes only by eliminating changes in price, by the use of the index of Wholesale Prices which was first introduced in 1948.[1]

Since it is the purpose of the index to compare the level of production in different months, corrections have to be made for the fact that calendar months do not all contain the same number of working days. Furthermore, some contain four, and others five Saturdays, a day on which production is likely to be lower than on the other days of the week. Such vagaries of the calendar have as far as possible to be eliminated. Most of the 1,300 series used in the index relate to the 'output' of weeks or calendar months. Table 66 below illustrates the nature of the data available in terms of the period to which it relates. It will be apparent from that table that some adjustment of the data relating to

[1] This index is described on pp. 400-3.

about one-third of the various series was necessary in the 1948 index. Since 1952, however, new data have become available and in the new index based upon 1954 the *annual* series shown in Table 66 have been replaced with short period statistics. No details have yet been published. One advantage of this change is that it will reduce the number of retrospective revisions that have had to be made in the past to published indices.

TABLE 66

Series classified by period covered by output data 1948.	No. of Series	Percentage of weight carried
Time Interval		
Calendar months	950	35·1
Weekly figures and weekly averages of periods of 4 and 5 weeks	135	27·0
Quarters	130	24·5
Yearly (without alternative series for shorter intervals)	40	2·7
Yearly (with alternative series for shorter intervals)	45	10·7
	1,300	100·0

Source: op. cit. H.M.S.O.

Apart from correcting the monthly indices for the varying number of working days in each month, separate indices are prepared which have been adjusted for holidays and other seasonal causes of variation in production. The object of this series is to eliminate the usual month-to-month fluctuations and to bring out the trend more clearly. Experience in this field has revealed that the seasonal pattern varies slightly from year to year, so the seasonal corrections applied to the monthly series need to be kept continuously under review and reassessed each year. The official account of the index emphasises that these seasonally corrected indices should not be regarded as in any way more reliable than the uncorrected series; nor are they intended to replace them. They emphasise the trend, whereas the original uncorrected monthly series measure the fluctuations in the weekly rate of production from month to month.

The weighting of the various indicators or series is based on the value of the net output of corresponding industries in 1954 as given by the Census covering that year. The net output of an industry, as defined for purposes of this Census, is the selling value at factor cost of services and finished and partly finished

goods produced by that industry during the year. Alternatively, it may be regarded as the gross output less the costs incurred in production of that output.[1] Certain adjustments have had to be made to the Census data, in particular with regard to the inclusion in Census aggregates of the amounts paid for services rendered by other industries. Allowance has also had to be made for the output of small establishments, since the Census figures of net output relate only to establishments employing more than ten persons. Even after deriving the net output each industry it was often necessary to apportion it over several products within the industry to provide appropriate weights for the individual indicators. In terms of index number theory, the index is the arithmetic mean of a collection of indices measuring the output of a large number of industrial goods each based upon the ratio of outputs between any given month and the base period, each index (or quantity relative) being weighted by reference to base period quantities (*i.e.* Census of Production net output) valued at 1954 prices.

The revised index which was introduced in 1958 differs from its predecessors, *i.e.* those based upon 1946 and later on 1948, in more respects than just its base year which is now 1954. The

TABLE 67

INDEX OF INDUSTRIAL PRODUCTION*

(Average 1954 = 100)

Weights	Standard Industrial Classification Order Number	Industry or industrial group	1955	1956	1957	2nd Quarter 1957	2nd Quarter 1958	1958 July
100 00	II-XVIII	All industries	105	106	107	108	106	98
7·24	II	Mining and quarrying	99	99	99	102	96	84
	III-XVI	Total manufacturing industries ..	106	106	108	109	107	98
8·18	III	Food, drink and tobacco	103	105	107	109	113	104
6·30	IV	Chemicals and allied industries ..	106	111	115	115	115	
6·87	V	Metal manufacture, Ferrous ..	108	111	113	115	102	82
	VI-IX	Engineering and allied industries ..	110	108	111	111	111	104
2·17	VII	Shipbuilding and marine engineering	108	117	108	105	108	107
7·78	VIII	Vehicles	115	107	115	115	120	116
	X-XIII	Textiles, leather and clothing ..	99	99	99	100	88	80
2·19	XIV	Timber, furniture, etc.	100	94	96	95	90	87
5·33	XV	Paper, printing and publishing ..	108	106	109	111	113	93
4·83	XVIII	Gas, electricity and water	105	110	112	103	108	96
		Seasonally corrected:						
	II-XVIII	All industries				108	106	106
	III-XVI	Total manufacturing industries ..				108	106	107

Source: Board of Trade Journal, 7th November 1958.

*Selected indices only.

[1] See pp 354-355 for definitions.

introduction of a revised Standard Industrial Classification in 1958 has led to a reclassification and regrouping of certain industries.[1] These changes, together with the new sources of data for the monthly indicators and revised weighting referred to above, effectively destroy the comparability of the different but successive series of index numbers. The Central Statistical Office has, however, calculated annual indices on the 1948 base using the revised Industrial classification so that it is possible to measure changes back to 1948 with the new index. Similarly it has calculated annual indices for the new index (1954 = 100) back to 1955. Actually the changes in both the weighting and revised industrial classification affect only the movements in the series for the individual industries; they hardly affect the index for 'all industries' which is the index usually quoted in economic discussion. Pending publication of a revised edition of the official account of this index, the student requiring more detailed information on these changes which are incorporated in the 1954 base index, is referred to the October 1958 issue of the *Monthly Digest of Statistics* and *Economic Trends* for November 1958.

While the substantial improvement in this index is freely acknowledged, it would be a mistake to place too much emphasis on month to month changes. For both the ordinary monthly and 'seasonally corrected' or adjusted indices it is the general persistent movement in one direction or another which is the most satisfactory and reliable guide to production levels.

Distribution

Almost a century and a half elapsed after Napoleon described this country as a 'nation of shopkeepers' before official action was taken to ascertain the truth of this comment! In 1939 nearly 15 per cent of the working population was engaged in the distributive trades, a total of 2·9 million persons. At the end of 1954 eleven per cent of the nation's total manpower was still so occupied. The term 'distribution' covers all the various channels through which goods pass from the manufacturer or grower (in the case of food) to the final consumer, *i.e.* all wholesalers and retailers. Included with these for the purpose of the Census are the service trades such as hairdressing, shoe repairing and garages; in other words, the section of the working community

[1] The changes in the S.I.C. are not major, but they are sufficient to affect the comparability of some of the individual industry or group indices.

not covered by the census of industrial production. The first census of distribution ever was taken in Great Britain by the Board of Trade in 1951. Before this, information on the extent of the distributive trades, the number of shops, scale of their activities as reflected in the number of employees, their wage bill and annual turnover were unknown. Even the number of distribution outlets could only be estimated at something over three-quarters of a million as compared with an actual figure of about 700,000.

A government committee set up in June 1945 recommended the taking of a census and the government acquired powers to conduct such an enquiry under the 1947 Act. The census was conducted by post during 1951 and the respondents were asked to provide information relating to their activities in 1950. The same concession regarding use of the firm's financial year in place of the calendar year was made as is given in the Census of Production. Traders were given three months to complete the forms but it may be noted in passing that not merely was the response very slow but that a number of prosecutions arose from failure to make the statutory return. Despite the great efforts made to ensure the co-operation of the traders, many were suspicious of the authorities' intentions and unwilling to co-operate. In consequence, the accuracy of some of the returns must inevitably be a matter for speculation.

Before the forms could be distributed it was necessary to carry out a census of the distributive and related service trade establishments in the country. This was done during May to October 1950; enumerators all over the country listed the names and addresses of traders apparently falling within the scope of the census – note – 'as far as could be judged from the outside of trading premises'. This is an interesting example of the difficulties encountered when a sampling frame is either non-existent as in this case and has to be built up, or is seriously defective. In this particular example the funds were made available for a complete enumeration before the census. The enumerators distinguished between shops, stalls, yards, depots and other types of premises. It should be noted that the enumeration staff had instructions not to enter the premises or question traders or their employees; the basis of their description as indicated above, was visual.

The purpose of the census – as distinct from the enumeration which preceded it – was to provide:

(1) information about the number and size of wholesale and retail outlets and other establishments providing consumer services.

(2) information regarding the value of the services rendered to enable more accurate estimates of the national product to be made.

(3) a measure of the relative efficiency of the distributive system as between different regions in the country, *e.g.* which areas have the most shops of certain kinds per head of the population and what is their turnover.

Apart from the above information which was primarily of interest to the government in respect of its economic policies, the census was to provide further information which would be of interest to traders and their trade associations. Quite apart from their natural concern with the distribution of various types of shops throughout the country, they would have an interest in the turnover, wage bills, level of stocks maintained and methods of delivery employed. Some of the data assembled to answer these questions are reproduced below from the official reports.

The data assembled as a result of the pre-census enumeration are reproduced in a publication entitled *Britain's Shops* and a detailed account is given of the enumeration, its difficulties and methods in the introduction to that report, which is better described by its sub-title, *A Statistical Summary of Shops and Service Establishments*. It contains a breakdown of the various retail shops into 22 classes, *i.e.* commodity traders, and service trade establishments into 5 classes. For each class of establishment the number in the country, in the counties, in the Metropolitan Boroughs and the City of London are given, as well as for towns of over 50,000 and 100,000 population respectively (excluding the metropolitan boroughs). For each class and area, the number of outlets per 10,000 population is given. This particular report is quite distinct from the Census of Distribution reports themselves. There are, of course, differences in classification and of coverage. For example, the census proper obtained a 91 per cent response (of all the outlets enumerated in the above enumeration) which was estimated to cover 95 per cent of the total trade of the retail establishments enumerated.

The Census of Distribution reports themselves are four in number; the three larger reports (one of which is devoted to Wholesale Trade) supplementing the Short Report which summarises the data collected for the retail trade. The results of the census describe the characteristics of retail trade in different parts of the country and in towns of various size. The data reveal the variations in gross margins (the best approximation to a measure of profit available from the data) between retailers in different lines of business as well as between retailers in the same line but of different sizes of establishment. Information relating to the distribution of sales of various kinds of goods between independent retailers, co-operative societies and multiple traders is also provided. Data showing the variation in stock levels and rate of stock turnover, wages and persons employed, as between different sized establishments are given. For the wholesale trade information is available as to the geographical distribution of these organisations, the main commodities handled and the distribution as between large and small wholesalers.

The data reproduced in Table 68 must be interpreted with due regard for the definitions employed in the Census, quite apart from the basic considerations already mentioned of incomplete coverage and the probability of inaccurate returns. A 'retail establishment' is defined as 'a separate place of business engaged in the sale of goods at retail prices'. The goods covered by the classification of shops employed in the report are set out in appendices to the reports. The establishment was assigned under a particular heading, *e.g.* Grocery, after due consideration of all the information available, *i.e.* enumerator's return, informant's own description, division of sales etc. The headings under which the establishments are classified are divided into two categories, 'specialist' and 'combined headings', *e.g.* grocers, and grocers with off-licence respectively.The breakdown of turnover was given for all establishments with an annual turnover exceeding £5,000. Multiple retailers and co-operative societies submitted analyses of their sales by commodity, not by branch.

Sales include all hire-purchase transactions originating in the year at their cash value plus any charge for credit provided by the retailer; the figures are shown inclusive of purchase tax where it applies. The data relating to employees are based on the return of all persons who worked in the business in the week

TABLE 68

RETAIL TRADE CLASSIFIED BY KIND OF BUSINESS. G.B. 1950

Kind of Business	Organisations	Establishments	Sales	Persons engaged	Gross Margin	Gross Margin as %age of Sales	Stocks end of year
	Number	Number	£000s	Number	£000s	%	£000s
Total in G.B.	403,839	528,450	4,930,176	2,378,635	1,092,249	22·2	571,103
Grocery group:	98,459	118,457	889,766	415,175	136,523	15·3	69,393
Grocers	68,603	85,417	680,777	305,142	100,028	14·7	54,942
„ with off licence	7,050	7,683	76,342	29,079	10,446	13·7	6,694
„ „ meat	1,902	2,681	28,497	12,197	5,721	20·1	700
„ „ bakery goods, with baking	5,330	6,187	40,338	32,427	11,461	28·4	2,328
„ „ „ „ without baking	13,972	14,557	53,005	30,787	7,087	13·4	3,296
„ „ hardware	1,602	1,932	10,807	5,543	1,779	16·5	1,433

Source: Census of Distribution 1950. Retail Trade. Short Report. Extract from Table 6.

ended 24th June 1950. Working proprietors and members of the owner's family are included if they worked in the shop. The 'gross margin' is the difference between the value of sales and the value of purchases for the year plus the net change in the value of stocks over the year. It represents, in effect, the retailer's margin out of which he must meet all business expenses including his remuneration.

The Verdon Smith Committee on the Censuses of Distribution and Production recommended in 1954 (Cmd. 9276) among other things, that full censuses of distribution should be taken every 10 years and that sample surveys should be taken from time to time in the interval between them. The first of these sample surveys was taken for 1957 and the preliminary results published early in 1959.[1]

One of the most interesting features of this survey was the manner in which the sample was selected. The aim was to include all traders with an annual turnover of over £100,000. These included all multiple retailers, all except the smallest cooperative societies, department stores, etc., but while it is fairly easy to enumerate all these larger traders with reasonable accuracy, there was no way of obtaining an up-to-date list of all other smaller independent traders. This object could only be achieved as in 1950 by enumerating them in a special census for that purpose. The remainder of the sample covering these independents was therefore taken on a geographical basis as follows:

(i) New Towns, Central London and a few special areas where great changes were thought to have occurred since 1950; all retail trades were enumerated and a sample of one in five taken.

(ii) Greater London: a sample of electoral wards stratified by size was taken, distinguishing between shopping and mainly residential areas, and all shops in the selected wards were included in the survey.

(iii) Large towns (population 100,000 or more): a sample of streets stratified by size, *i.e.* number of shops in 1950, was taken.

(iv) Other towns were sampled by taking a cross-section of the local authority areas and sometimes stratifying the areas by

[1] Special Supplement to *Board of Trade Journal*, 2nd January, 1959.

size, by sales in 1950, and by population change since 1950.

(v) Rural Districts were also sampled by region after stratifying by size, either by population density or by population change since 1950.

The census was carried out by post with a very energetic follow-up which gave an 89 per cent return from independent retailers, compared with a 96 per cent response from the larger traders, *i.e.* multiples, etc. So thorough was the follow-up that returns were obtained from 75 per cent of street traders, pedlars, hawkers and itinerant market traders![1] The 1957 totals were estimated by compiling 1950 as well as 1957 figures for the sample and calculating the ratio of 1957 to 1950 figures. These ratios were applied to the known 1950 totals to give the 1957 estimates.

The information obtained in the census included the number of establishments, turnover and number of people engaged in the establishments. All this information is analysed both by the form of the organisation and the kind of business. An example of the former is given in Table 69. This table gives a breakdown by turnover and the number of establishments for three major forms of retail organisation, *i.e.* reading across the table – cooperatives, multiples, etc., and by the type of trade they carry on, *i.e.*, reading downward. From the point of view of tabulation, one comment might be made, *i.e.* the figures of turnover could better have been given to the nearest £ million instead of £000 as in the excerpt reproduced on page 371.

Since the 1957 survey results are partly based upon sample data, the published statistics are subject to sampling errors. These have been calculated for each figure under the general classification of 'independent traders', *i.e.* Turnover, Number of establishments and persons engaged, for each category of business of trade. Allowance for the effect of the error on the total results, *i.e.* independent plus large-scale retailers, is also made in the corresponding figures for 'all traders'. However, for all the industry an error of less than half of one per cent is given. The error is larger than this for individual items but considering that the sample was a 12 per cent one, *i.e.* covering about 57,000 establishments out of a total of 480,000, the accuracy seems

[1] A full account of the tabulating and estimation processes is given in the *Board of Trade Journal* Supplement *op. cit.*

TABLE 69

Retail Establishments Grouped by Form of Organisation and Kind of Business

	All Retail Establishments				Retail Establishments of Co-operative Societies				Retail Establishments of Organisations other than Co-operative Societies having 1 to 9 Branches				Retail Establishments of Organisations other than Co-operative Societies having 10 or more Branches			
	1950		1957		1950		1957		1950		1957		1950		1957	
	Number 000's	Turnover £m.	Number 000's	Turnover £m.	Number 000's	Turnover £m.	Number 000's	Turnover £m.	Number 000's	Turnover £m.	Number 000's	Turnover £m.	Number 000's	Turnover £m.	Number 000's	Turnover £m.
Total Retail Trade ..	580	5,100	573	7,790	26	572	29	905	497	3,377	482	4,908	57	1,151	62	1,977
1. Grocers and Provision Dealers ..	141	1,232	149	2,040	11	284	13	463	114	689	120	1,104	16	259	16	473
2. Other Food Retailers	144	1,006	127	1,560	9	148	10	244	123	693	104	1,037	12	165	12	278
3. Confectioners, Newsagents & Tobacconists	75	503	77	706	0·2	3	0·2	4	69	443	71	616	5	57	6	85
4. Clothing & Footwear	96	932	93	1,150	3	56	3·0	56	81	578	77	633	12	298	13	460
5. Household Goods ..	65	571	65	884	0·8	21	1·0	28	58	412	56	635	6	131	8	221
6. Other Non-Food Retailers	58	385	58	579	1·1	10	1·0	18	52	286	52	421	5	89	5	140
7. General Stores ..	2	471	4	872	0·2	50	0·7	92	0·4	269	11	461	1	152	2	319

Source. Board of Trade Journal, 2nd January 1959.

reasonable and quite adequate for administrative and statistical needs.[1]

REFERENCES

Report of the Census of Distribution Committee, Cmd 6764, H.M.S.O. 1946.
Britain's Shops, H.M.S.O. 1952.
Census of Distribution Reports, H.M.S.O. 1953.
The Census of Distribution, J. I. Mason, The Incorporated Statistician, August 1953.
Report of the Committee on the Censuses of Distribution and Production. Cmd. 9276. H.M.S.O. 1954.

Retail Sales Indices

For over twenty-five years there have appeared monthly in the *Board of Trade Journal* indices relating to the level of retail trade. The indices have been subject to several revisions, in particular in 1947, 1952 and 1955. In 1952 the base year was changed to the current one of 1950.[2] In 1955 the basis of presenting the statistics was changed from a *commodity* basis, *i.e.*, showing sales of furniture, sales of groceries, etc., to a *kind of business* basis, showing sales by furniture shops, sales by grocers, etc.; the base year remained 1950.

The whole field of retail trade in Great Britain is covered, with only a few exceptions such as coal merchants and florists. Retailers are divided into four groups: independent retailers, multiple retailers (chains of ten or more branches), cooperative societies, and departmental stores. The basis for comparison for each of the categories mentioned above is the sales as shown in the Census of Distribution of 1950.

The currently published statistics are based on voluntary returns covering some 7,500 shops owned by independent retailers and about 60 per cent of the sales of all large-scale retailers. It would be too large a claim to assert that these form a random sample. Attempts are being made to improve the representativeness of the returns. Prior to 1955 these statistics were shown as indices calculated on a chain basis, but following the 1955 revision the indices are compiled using a ratio method. The ratio method consists of dividing the total sales of all contributors to the index in any one category in the current period by their total sales in 1950, multiplying the result by the total sales of all shops in that category in 1950. Combined index numbers are calculated in a similar way. Thus:

[1] See Board of Trade Supplement *op. cit.* for comments.
[2] Base changed again this year to 1957 = 100; but the structure of the indices remains unchanged.

$$\text{Index of the sales in January 1959} = \frac{\text{Sales of the sample in January 1959}}{\text{Sales of the sample in 1950}} \times \text{Sales of all shops in 1950}$$

Table 70 below gives an example of how the main body of statistics are presented. Similar detail is given for each kind of business classification used in the Census of Distribution. In addition to the main 'kind of business' statistics, the returns of departmental stores are analysed into commodity statistics. These show the percentage changes in values of weekly sales and

TABLE 70

INDICES OF VALUES OF SALES PER WEEK AND PERCENTAGE CHANGES COMPARED WITH A YEAR EARLIER. 1950=100

		1957	1957		1958			1957			1958		
		Year	3rd Qtr.	4th Qtr	1st Qtr.	2nd Qtr.	3rd Qtr.	July	Aug.	Sept.	July	Aug.	Sept.
Total Food Shops	Total All Retailers	163	162	173	164	166	164	164	161	160	167	162	163
		+5	+5	+4	+4	+3	+1	+7	+4	+5	+2	+1	+2
	Independent Retailers	157	156	166	155	158	158	159	157	153	162	157	155
		+3	+4	+3	+3	+2	+1	+5	+2	+4	+2	0	+1
	Multiple Retailers	177	173	192	179	184	181	176	172	172	185	177	181
		+7	+8	+7	+7	+5	+5	+10	+6	+7	+5	+3	+5
	Cooperative Societies	169	165	176	173	171	165	166	160	168	165	161	170
		+6	+6	+4	+4	+1	0	+8	+6	+5	—1	0	+1

Source: Board of Trade Journal, various dates.

stocks compared with a year earlier. The changes in weekly sales are converted into indices by the method of changing back to the corresponding month a year earlier.[1] As with so many economic statistics the accuracy of some of these group data is dubious. More reliance can be placed on the trend, or any marked interruption in the trend, than in the absolute movements themselves.

Hire Purchase Statistics

Since 1945 hire purchase has come to play an increasingly important part in the national economy. By 1959, credit outstanding amounted to nearly £600m. In consequence of this rapid expansion and its impact on the economy, it has become necessary to find some measure of the changes in both the volume and character of hire purchase debt. The *Board of Trade Journal* is publishing statistics monthly which cover the period since October 1955. These relate to the hire purchase trade of retailers in kinds of businesses where substantial sales of goods on hire purchase terms are made (*e.g.* furniture, radio and

[1] A good account of these data is given in the *B. of T. Journal* for May 5th, 1956. See also *B. of T. Journal*, February 6th 1959, for revision to current basis.

electrical shops) as well as to the hire purchase business of finance houses.

The information is collected on an extensive sample basis similar to that used for obtaining the figures of retail sales. In this case about 1,600 independent retailers make returns. They are divided into some 700 furniture and furnishing shops and 900 radio, electrical and hardware shops. The department store sample alone accounts for about a third of the total turnover of goods usually bought on hire purchase, while the multiple organisation return covers about half the turnover in furniture. The Cooperative Union returns details of over half the hire purchase business done by cooperative societies. Information is also collected from the area boards of the nationalised gas and electricity undertakings. About 260 finance houses contribute to the Board of Trade's information on direct credit retailing.

Data are collected on a voluntary basis and are subject to possible errors of bias, *e.g.* the sample is not random; some shops with a lot of hire purchase business may not bother to make a return. Consequently too much significance must not be attached to any particular figure but generally the trends in the volume of business are fairly indicated.

The information is presented in five tables which show:

(1) estimated total outstanding hire purchase debt;
(2) indices of value of goods sold on hire purchase by household goods shops;
(3) percentage of hire purchase to total sales by household goods shops;
(4) index of new hire purchase extended direct to hirers by finance houses;
(5) average value per agreement of new hire purchase by finance houses.

Explanatory notes are appended to the published tables.

The two *indices* measure changes in value of new hire purchase business, the first at household goods shops broken down by type of shop, and the second by finance houses broken down by type of goods. They are now based upon July 1957 = 100. Extracts from tables of both indices are given below. For the first two main categories of shop in Table 71 there is given not only a total index, but separate indices for sectors of each main

category of retail outlet, *e.g.*, multiple and independent radio and electrical goods shops. The content of Table 72 requires no comment. Note incidentally, the seasonal character of much H.P. business as reflected in the indices.

The original base of both indices was December 1955 = 100.

TABLE 71

INDEX OF VALUE OF GOODS SOLD ON HIRE PURCHASE BY HOUSEHOLD GOODS SHOPS

July 1957 = 100

		1957	1958		
		Oct.	Jan.	April	Oct.
FURNISHING AND FURNITURE SHOPS	Total, all classes of shops of which	120	92	94	164
	(*a*) Multiple Retailers	119	92	96	166
	(*b*) Independent Retailers	121	94	91	158
HARDWARE, RADIO, ELECTRICAL GOODS, CYCLE AND PERAMBULATOR SHOPS	Total, all classes of shops of which: (Multiple and independent):	129	112	91	151
	(*a*) Radio and electrical shops	175	149	104	169
	(*b*) Cycle and perambulator shops ..	98	104	71	89
DEPARTMENT STORES: Household goods department		132	103	98	179
TOTAL, HOUSEHOLD GOODS SHOPS.. ..		124	101	93	159

Source: Board of Trade Journal, 23rd January 1959.

TABLE 72

INDEX OF VALUE OF NEW HIRE PURCHASE EXTENDED DIRECT TO HIRERS BY FINANCE HOUSES*

July 1957 = 100

	1957	1958		
	Oct.	Jan.	April	Oct.
Private cars – new	79	82	99	58
Private cars – used	78	75	108	89
Motor cycles, side cars etc. – new and used..	67	49	96	114
Farm equipment and tractors	86	55	84	100
Industrial and building plant and equipment	121	106	118	142
Furniture, furnishings and floor coverings..	129	112	104	175
Domestic appliances	84	73	87	127
All goods (including those not shown above)	85	84	101	98

Source: Board of Trade Journal, 23rd January, 1959.

*Selected categories of goods only are shown in this table.

The changes introduced when the base was revised to July 1957 = 100 made comparison between the new and old based figures difficult, although the changes were not large enough to alter the character of the statistics entirely. In consequence, general trends since 1955 can still be traced with some confidence.

National Income

The national income may be defined as the money value of the nation's output of all goods and services in a given period, usually a year. This aggregate is also referred to as the national output, since these incomes represent the cost of producing the output of goods and services. It follows, therefore, that there are two ways of measuring the national income; either all the incomes of the factors of production, or the values of each industry's output, may be aggregated. Before discussing the various problems that arise in measuring these aggregates, it is important to understand their purpose. It is clearly desirable to know what the nation's economy is producing in any year, as well as comparing one year with another to ascertain the rate of economic progress. All production is intended ultimately to satisfy consumer needs. The more that is produced, the more the community may consume, *i.e.*, the higher its standard of living. Finally, the national income estimates are so prepared and presented, that they offer a comprehensive picture of the operations of the economy and the inter-relationships between various sectors. To the extent that the statistical data assembled in the annual Blue Book on the National Income and Expenditure are complete and accurate, overall economic planning is greatly facilitated.

National income estimates were first prepared officially by the Treasury in 1941, and used by the Chancellor as a background to the Budget statement in that year. In each of the next ten years a White Paper on these estimates was published a short while before the Budget statement. Since 1952 a Blue Book, which appears in autumn, has appeared annually. The pre-Budget document is a short White Paper containing preliminary estimates of the main aggregates. Before 1941 there had been several private estimates of the National Income or Output. Various methods were used to arrive at these estimates. Unfortunately the accuracy of the figures was greatly impaired by

the shortcomings of the data then available. Despite great improvements, even now the limitations of the data are such that in each year the successive Blue Books contain amended figures for previous years. The volume of data on which these estimates may be based has been considerably expanded in recent years, but even eighteen years after the first paper appeared, several of the more important aggregates remain little more than approximations.

The published data are based upon material derived from three main sources, although these must be supplemented by information culled from a wide range of other sources. Even so the coverage is in many cases incomplete, while further difficulties arise from the fact that much of the published data used has in fact been compiled for purposes other than national income estimates.

The three main sources of data are the statistics assembled by the Inland Revenue, the censuses of production and distribution and lastly the accounts of the central government. The significance of the last mentioned source may be better appreciated when it is remembered that the government is responsible for the expenditure of about one-third of the national income. Of these data those derived from the Inland Revenue are the most complete and accurate; those compiled from the Census of Production the least reliable.

The national income can be visualised in three ways:

(i) as a sum of incomes derived from economic activity, *i.e.*, from employment and profits;

(ii) as a sum of expenditure, *i.e.*, consumption and investment;

(iii) as a sum of the net products of the various industries of the nation.

These three views of the national income tend to explain the ways in which the statistics are presented and the estimates compiled. Some of the more frequently employed aggregates are given in Table 73 below, which illustrates the income approach in practice, *i.e.*, those countries with a well-developed fiscal system. The various types of income are given in the upper part of the table and they are largely self-explanatory. The *residual error* is the balancing figure between the two separate estimates of the gross national product, the one based on incomes and the other on expenditure. The sub-aggregate is described as the total

TABLE 73

GROSS NATIONAL INCOME ANALYSED BY FACTORS. SELECTED YEARS 1948-1957

Factor incomes	All figures £'s millions			
	1948	1951	1954	1957
Income from employment	6,766	8,459	10,263	12,942
Income from self-employment ..	1,320	1,450	1,588	1,787
Gross trading profits of companies ..	1,798	2,489	2,603	3,265
Gross trading surpluses of public corporations	118	258	348	333
Gross profits of other public enterprises	106	120	111	131
Rent	419	511	725	862
Residual error	8	84	137	— 68
Total domestic income before providing for depreciation and stock appreciation	10,535	13,371	15,775	19,252
Stock appreciation	— 325	— 750	— 75	— 100
Gross domestic product at factor cost ..	10,210	12,621	15,700	19,152
Net income from abroad	187	217	228	226
Gross national product	10,397	12,838	15,928	19,378
Capital consumption	— 890	—1,146	—1,424	1,774
National income	9,507	11,692	14,504	17,604

Source: National Income and Expenditure 1958. Table 1.

domestic income before depreciation and stock appreciation. These items inflate all the above incomes except rent and income from employment. An adjustment is made to eliminate the element of stock appreciation which may be defined as the increase in money terms in the value of stock distinct from a change in its physical quantity. The figures given for this item are little better than guesses, 'hazardous approximations' is the official description. Nevertheless, as will be seen from Table 73 it is an extremely important item, more especially in periods of rapidly changing prices, *e.g.*, as in 1951.

The figure described as the *gross national product* at factor cost should be distinguished from the total defined as the net *national income*. The difference between the net and gross figures is accounted for by the depreciation of capital equipment in the country. Unfortunately the data relating to depreciation or 'capital consumption' are extremely unreliable and incomplete. Rather than guess at this figure it has been considered better to omit it from the published estimates in some years. As an offset to this omission the 1956 and later Blue Books contain extended tables of investment and its consumption. The 1958 Blue Book

gives a detailed account of the sources and methods used to arrive at these estimates. These data are derived from a variety of sources; *e.g.*, local authority accounts to cover housing outlays; motor vehicle registrations to cover investment in road transport, the published accounts of public corporations and public companies as well as the statistics of stocks of primary commodities. These data are by no means comprehensive and the gaps to be filled are many.

The second method of estimating the gross national income is given in Table 74. The correspondence between the value of the national product and national income was mentioned earlier. In this table the gross products of various industries and sectors of the economy are given for four years. Similar adjustments in respect of stock appreciation and the residual error are made in this table. Most of the data upon which these figures are based are derived from the censuses of distribution and production, but since these are held only at intervals, the reliability of the

TABLE 74

GROSS NATIONAL PRODUCT BY INDUSTRY OF ORIGIN

Industry	All figures £ millions			
	1948	1951	1954	1957
Agriculture, forestry, fishing	644	716	761	850
Mining and quarrying	384	447	558	702
Manufacturing	3,739	4,961	5,915	7,279
Building and contracting	571	677	903	1,121
Gas, electricity and water	210	273	367	484
Transport and communication	880	1,152	1,267	1,659
Distributive trades	1,393	1,751	1,939	2,383
Insurance, banking and finance	281	372	459	566
Other services	1,017	1,188	1,283	1,647
Total production and trade	9,119	11,537	13,452	16,691
Public administration and defence	670	822	983	1,166
Public health and educational services	260	402	505	665
Ownership of dwellings	296	367	532	616
Domestic services to households	110	95	95	97
Services to private non-profit-making bodies	72	64	71	85
less Stock appreciation	— 325	— 750	— 75	— 100
Residual error	8	84	137	— 68
Gross domestic product at factor cost	10,210	12,621	15,700	19,152
Net income from abroad	187	217	228	226
Gross national product	10,397	12,838	15,928	19,378

Source: National Income and Expenditure 1958. Table 10.

final results cannot be as great as could be wished. By carrying out sample surveys in the fields of both industry and distribution in the years between the full-scale censuses, the Blue Book estimates for these years are much improved.

Table 75 shows the gross national product by categories of expenditure. This particular method remains the least satisfactory of the three, particularly in this country where data relating to consumer outlays on various commodities and services are extremely unreliable and subject to a considerable margin of error. The composition of the aggregate in this table requires no explanation, except to point out that taxes on expenditure are the outlay taxes which inflate the prices of goods which are purchased by consumers and public authorities, hence the sub-total of £21,609m defined as 'expenditure at market prices'. To change 'market prices' to 'factor prices' outlay taxes must be deducted and subsidies added back. In connection with these data, it should be noted that the Blue Books contain detailed analyses of consumer expenditure over a period of years. To bring out more clearly the shifts in consumer outlays between different categories of goods and services, the annual money outlays are corrected for price changes.

The figure of stock appreciation which appears in these tables came into prominence owing to the very sharp rise in commodity prices during 1951 as a result of the Korean War.

TABLE 75

GROSS NATIONAL PRODUCT BY CATEGORIES OF EXPENDITURE

(£ millions)

	1948	1951	1954	1957
Consumers' expenditure	8,475	10,085	11,984	14,174
Public authorities' current expenditure..	1,762	2,443	3,139	3,583
Gross fixed domestic capital formation..	1,455	1,921	2,583	3,402
Value of physical increase in stocks and work in progress	175	575	50	450
Total domestic expenditure at market prices	11,867	15,024	17,756	21,609
Exports and incomes from abroad ..	2,392	4,008	4,207	5,244
less Imports and incomes paid abroad ..	—2,412	—4,388	—3,972	—4,932
less Taxes on expenditure	—2,023	—2,274	—2,486	—2,956
Subsidies	573	468	423	413
Gross national product at factor cost ..	10,397	12,838	15,928	19,378

Source: National Income and Expenditure 1958. Table 1.

Note that in Table 75 there is no such figure; only the *real* not money change in stocks is included. The profits arising from the revised valuation of stocks held by companies and trading concerns inflate the annual trading profits for the relevant years, and tax was assessed on these profits. It is a moot point whether such 'income' should be included, but in view of other arbitrary decisions that have to be made in computing the national income, its inclusion in most years makes no significant difference. An important omission, however, is the value of housewives' services in the home which are unpaid. If they were valued in money terms, the national income in money terms would rise by at least 20 per cent. Colin Clark has attempted a more detailed estimate of the value of housewives' services. A full account of this estimate is published in the Bulletin of the Oxford Institute of Statistics.[1] Alternatively, if the housewives went out to work and paid some of the present factory or office workers to do their housework, although the real national income would not have changed, it would have increased considerably in money terms. Much the same conceptual problem arises with work for which no payment is received. Thus, a man working an allotment and consuming the produce at home adds nothing to the 'national income'. If, however, he and his neighbour sold each other their respective produce, the value of the produce added should, in theory, be added into the national income. Other problems arise in connection with the services of government, *e.g.*, should the salaries of Civil Servants be included since they add nothing tangible to the national product? The same argument can perhaps be applied with more justice to the payments to members of the armed forces in peace-time.

These conceptual differences as to what should be included in the 'national income'make international comparisons extremely difficult, *e.g.*, valuation of home-produced food in an agrarian economy. There have been several conferences under the auspices of U.N.O., with the purpose of standardising practice. Apart from these problems the differences in the reliability of various statistical data in different countries pose a problem which will probably not be solved for many years yet. Any published international comparisons should be scrutinised with these considerations in mind.

[1] Vol. 20. No. 2. May 1958.

The advantages of computing the national income totals by various methods are obvious. The numerous cross-checks which are thereby made possible, especially in the various sub-totals, are invaluable. The modern method of constructing the National Income accounts, known as 'Social Accounting', is simply the adaptation of double-entry book-keeping principles.[1] Its value lies in the fact that every sub-total appears twice in the accounts and if it does not 'fit' in with the expected value as indicated by the size of the other totals in that particular section of the accounts, there is presumably some error. The difficulty is that a system of statistics of national income and expenditure must be comprehensive to be of use and an estimate must be included for each item that appears in a balancing account. It is not possible to base all the estimates on accurately recorded facts nor is it possible to calculate statistical 'margins of error' of the kind derived from random samples. What is done, however, is to form very rough judgements of the range of reasonable doubt attaching to the estimates. Some standardisation is obtained by grading each of the major components as having a margin of error of: A $\pm$ less than 3 per cent,
B $\pm$ 3 per cent to 10 per cent,
or C $\pm$ more than 10 per cent.

As far as the various methods of deriving the various aggregates are concerned, the correspondence of the three aggregates of the national income, product and outlay, while not necessarily proving the accuracy of any one of them, would suggest that the errors are not such as to invalidate the overall results. It is certain, however, that the major aggregates are much more reliable than the numerous sub-totals in the analyses. There still exist several important gaps in the requisite information for any one of the three approaches. For example, wages in the income approach, distribution in the output data and items such as motoring and holidays in the analysis of current personal expenditure, must be interpreted with caution. The outstanding weaknesses in the over-all aggregates remain the deficiencies in the data from which estimates of the level of savings and net investment in this country can be made. The Central Statistical Office has completed two major enquiries into the current levels of savings and investment which have been read before the two

[1] For an account of the construction of such accounts see 'National Income and Social Accounting' by Edey and Peacock, Hutchinson.

statistical societies.[1] Although these researches are much more thorough and detailed than anything done hitherto, there are still gaps which cannot be filled. Until more detailed information regarding not merely the volume of savings and investment, but in particular their distribution within the economy is known, economic planning must remain a highly speculative exercise. Nevertheless, despite all the criticisms of the data as published, it is no exaggeration to state that the Blue Book on the National Income and Expenditure is the most important economic document of the year. The Economic Survey is based on that information and estimates. Without them fiscal and budgetary policy would be mere guess work. As the volume of statistical data assembled by the government increases, as it undoubtedly will, so the accuracy of these estimates will be improved.

The Central Statistical Office published in 1956 a very full account of the sources and construction of the statistics of National Income entitled 'National Income Statistics: Sources and Methods'. The first three chapters are probably the most useful for students as the remainder of the book goes into great detail. It is, however, an excellent source of reference. The notes which accompany the Blue Book on National Income consist:

(1) of definitions of items in the summary tables in the Blue Book;
(2) of revisions made in the previous years' estimates; and
(3) changes in treatment and definition made since the publication of the Central Statistical Office study mentioned above.

REFERENCES

Annual White Papers and Blue Books on National Income and Expenditure.

'Use and Development of National Income and Expenditure Estimates' by R. Stone in *Lessons of the British War Economy*, Ed. D. N. Chester.

British Economic Statistics by C. F. Carter & A. D. Roy, Chapter IX.

'National Income Statistics: Are they accurate or useful?' C. T. Saunders in *The Incorporated Statistician*, March 1954.

'National Income and Related Statistics.' J. E. G. Utting. J. R. S. S. 1955, Pt. IV.

Overseas Trade

Statistics of overseas trade are among the oldest to be prepared in the U.K. They date from the establishment in 1696 of

[1] 'The Pattern of Savings and Investment' by C T. Saunders. Transactions of the Manchester Statistical Society 1954, and 'Net Investment in Fixed Assets in the U.K. 1938-1953' by Philip Redfern, J.R.S.S. Vol.118 Pt.2, 1955.

the office of 'The Inspector-General of the Imports and Exports'. Their origin was to be found in the need to collect revenue and even today, despite the considerable changes and improvements, the classification still bears traces of this purpose. The statistics as compiled at the present time effectively start in 1871 when the statistics were based upon importer's and exporter's declarations of value (as well as of quantity) collected by Customs Officers at the ports and transmitted to the Customs Statistical Office for compilation. It is of some interest to note that since 1871 these data have been affected by two major changes only; the inclusion of exports of ships and boats as from 1899 which then represented about $3\frac{1}{2}$ per cent of the total value of exports and again in 1923 when the Irish Free State was created and trade with that country then became part of the U.K. external trade. The only other changes concerned classification, in particular, of countries due to changes in their frontiers.

The statistics of U.K. overseas trade are based upon the official certificates or declarations which must be made by both importers and exporters. These certificates give details of the nature of the merchandise together with figures of quantity and value. Imports are valued c.i.f. and exports and re-exports are valued f.o.b. The first abbreviation means 'carriage, insurance and freight'; the practice of including these items with the cost of the commodity imported follows logically from the definition of the value of imports required by the Customs; *i.e.* the 'open-market' value or price inclusive of all costs of importation, which the merchandise would fetch if sold on the open market at time of entry into this country. The valuation placed upon exported merchandise represents the cost of the goods packed and delivered to the ship, *i.e.* f.o.b. The c.i.f. basis of valuing imports is something of an anachronism; it started as a result of the 1932 Import Duties Act which created a general *ad valorem* tariff *i.e.* a tax on imports based upon the value thereof and clearly some standard method of valuation was needed. Actually this method did not differ greatly from the mode of valuation employed up till 1932, which was the cost to the importer including freight, insurance etc. Nowadays about 85 per cent of the imports are duty free and the need for precise valuation in accordance with the formula is weakened. In practice, it is the cost price, *i.e.* the price actually paid for the goods to the port

of entry, which the importer records on his certificate. This may differ substantially from the 'open-value' price if for example, as occurred in 1949, a devaluation takes place and goods are imported at their cost before the devaluation occurred. For the dutiable goods, the authorities are satisfied that the 'total values reflect fairly accurately the actual c.i.f. cost of imports'.[1]

The statistics are first published in the monthly Trade and Navigation Accounts which appear about the twentieth day following the end of the month covered by the Accounts. The monthly Accounts cover imports, exports and re-exports; the latter comprising that merchandise which is exported in virtually the same form as it was imported. The classification is extremely detailed, with about 800 headings for imports and some 1,200 for exports. Details of the country of origin or destination for imports and exports respectively are given for the main headings. In addition to the monthly figures, the Accounts contain cumulative figures for the expired part of the year, *e.g.* the June accounts give the first half-year's aggregate, and these data for the current year are repeated for the preceding year for comparative purposes. The tables are preceded by a short introductory Note; any changes in classification are always announced in these Notes. These monthly tables are extremely detailed and not suitable for all requirements. Since March 1950 the Board of Trade has produced a condensed monthly Report on Overseas Trade based upon the monthly Accounts which re-classifies and condenses the original data into more useful form for general purposes, *e.g.* commodity analysis by different trading areas. These are supplemented at quarterly intervals by statements and tables on the nation's trade, *i.e.* value and volume in the main categories of merchandise, which appear in the *Board of Trade Journal.* These are accompanied by the indices of prices and volume for both imports and exports discussed below.

The companion volumes of the monthly Accounts are the four massive volumes which comprise the Annual Statement of Trade which appears usually more than a year after the end of the relevant year. The amount of information included in these volumes is such that 'the difficulty in use may be because of an excess of

[1] International Trade Statistics by R.G.D. Allen and J. E. Eley. Section on U.K. statistics prepared by J. Stafford, J. M. Maton and M. Venning p.302.

detail rather than a lack of it'.[1] Each volume contains comparable data for the preceding four years. The first volume provides an analysis of the value and volume of goods traded for each commodity or product. The second volume is concerned only with imports and re-exports, classifying them by country of origin and for each commodity giving value and quantity imported. Volume 3 is identical with the second, except that it analyses exports by country of destination. The last of the four volumes summarises the details of goods traded and gives a detailed country analysis of U.K. trade; *e.g.* imports and exports by value and volume for Sweden, Switzerland, etc.

The classification of both imports and exports is regularly brought up to date, the list of headings being revised annually; 'a balance being kept as far as possible between the need for continuity and for an up-to-date system, and between the need for detail and what it is practical to require from traders.'[2] On the import side, foodstuffs and raw materials form the main bulk of the trade, but on the export side, a list of 2,500 headings of which over 2,000 relate to manufactured goods, offers scope for inaccurate classification by the trader.

The main commodity headings under which U.K. imports and exports are classified were revised as from 1954 to cor-

TABLE 76

VALUE OF U.K. IMPORTS AND EXPORTS 1957 AND 1ST QUARTER 1958

Class and Division	Exports		Imports	
	1957	1958 1st Qtr.	1957	1958 1st Qtr.
	£000	£000	£000	£000
A. Food, beverages and tobacco	206,196	44,354	1,496,441	353,235
B. Basic materials	122,986	28,292	1,169,361	241,403
C. Mineral fuels and lubricants	152,704	38,213	466,302	104,469
D. Manufactured goods ..	2,754,375	690,703	928,315	230,767
E. Postal packages	82,716	18,073	7,838	2,543
Live animals not normally used for food	6,003	1,380	7,331	1,260
Total U.K. exports and imports	3,324,981	821,015	4,075,588	933,677

Source: Report on Overseas Trade, Vol. IX, No.7.

[1] J. Stafford and other *op. cit.* p.306.

[2] J. Stafford and others *op. cit.* p.298.

respond with the Standard International Trade Classification the basis of which is shown in Table 76.

In the case of both imports and exports the bulk of the trade is clearly concentrated into the first four classes. For exports the fourth class is by far and away the largest, for imports the first and second predominate. The classes are further subdivided: Class A into 10 divisions, B into 12 and C and E 2 each, while D contains 23 divisions. Thus division D 16 is 'electric machinery, apparatus and appliances. The description 'Postal packages' applied to the penultimate category in Table 76 is misleading. All the contents of parcels liable to duty are actually classified under the appropriate commodity heading and the figure entered as Parcel Post is simply an estimate based upon the product of the number of parcels containing non-dutiable goods and an average parcel value. The information upon which this estimate is based is derived from the Customs declaration which accompanies every parcel imported or sent overseas.

In passing, it should be noted that details of volume and quantity are available for about 98 per cent of all imports (by value) and 90 per cent of exports. This is important since such data enable index numbers to be calculated which permit the money aggregates over a period of years to be adjusted for changes in both price and quantity. The difficulties arising from changes in quality and type of product cannot be completely overcome by an index, hence it is sometimes difficult to be sure that one is comparing like with like. For example, the value of machinery which is adjusted by reference to its weight, will be affected by the growing use in recent years of lighter alloys for its construction.

So brief an account of the statistics of overseas trade can do little more than warn the reader who anticipates consulting any of the references given that the utmost care in extraction of figures is necessary. Comparability over a period of years is often more apparent than real and the notes to the various Tables and Accounts must be examined for changes, especially in classification. The difficulties of the student are intensified by the importance of the Balance of Payments problem; it is tempting to regard the statistics of overseas trade as a means of interpreting the balance of payments. This is far from being the case; the government issues a half-yearly White Paper on the balance of payments which is very different from the publications discussed

above. In fact, not even the expert can reconcile the documents since they are compiled on different bases. A major difficulty, to which reference is sometimes made in the press discussions of monthly trade accounts, is the problem of adjusting the cost of imports from the c.i.f. valuation to that for the goods themselves, *i.e.* ex insurance and transport costs. These costs are estimated to represent between 10 and 13 per cent of the total c.i.f. value, but the percentage varies as between the various commodities and as between different countries. At best the correction can only be approximate.

Import and Export Price Indices

Since 1946 there has appeared monthly in the *Board of Trade Journal* a series of monthly indices which are used to measure the short-period changes in the U.K.'s terms of trade. The 'terms of trade' is simply the ratio of import to export prices; if the former are rising more rapidly than the latter, then the terms of trade are said to be moving against the U.K. In other words, the U.K. is receiving a smaller quantity of imports for a given volume of exports. In view of the nation's post-war balance of payments problem this particular index is quoted regularly in discussions of the overseas trade statistics.

The price indices are designed to measure the monthly changes in the aggregate value of a *fixed* but representative selection of imports or exports. The index for export prices is based upon data relating to some 280 items, and that for import prices on about 220. In terms of the value of goods traded they cover about three-quarters, the proportion of total goods traded represented by the selected items being rather greater for the import index than for the export index. The indices are based not upon the actual prices of the goods included, but upon 'unit-values'. These are specially computed 'prices' of groups of homogeneous commodities which are classified under separate headings in the monthly Trade Accounts. The classification in the monthly Accounts must be broken down into greater detail for purposes of this index. In computing the unit values it is necessary to ensure that the prices of the constituent items do in fact move together. If any particular price in any month is erratic to such an extent that in that month the balance of the index would

TABLE 77*

IMPORT AND EXPORT PRICE INDEX NUMBERS

(1954 = 100)

Period	Imports					Exports						Terms of Trade
		(Class A)	(Class B)	(Class C)	(Class D)		Manufactured Goods (Class D)					
	Total	Food, Beverages and Tobacco	Basic Materials	Fuels	Manu-factured Goods	Total	Total	‡ Metals	‡ Engineering Products	‡ Textiles (excluding clothing)	Other	†
1950	85	84	83	101	82	85	85	81	82	90	88	100
1953	101	97	103	108	105		101	105	100	99	101	100
1955	103	101	104	102	109	102	102	104	103	100	102	101
1956	105	101	106	110	110	106	106	112	107	99	104	99
1957	107	101	110	128	104	111	109	115	112	101	106	96
1958	99	97	96	112	101	110	111	118	115	101	105	90
1958 Oct.	98	98	92	110	102	109	111	118	116	98	104	90
Nov.	99	100	91	111	104	109	111	117	116	97	104	91
Dec.	99	100	90	110	104	110	111	117	117	98	106	90
1958 4th Qtr	99	99	91	110	103	109	111	117	116	98	105	91

*Source: *Board of Trade Journal*, 13th and 27th February, 1959.

† Import price index as a percentage of the export price index. A rise indicates an adverse movement.

‡ These are defined as in the Trade Accounts classification.

be upset, the distorted value may be replaced by a 'smoothed' figure.

The selection of the unit values under each heading is so devised that a collection of commodities for both indices is derived which is representative of the current pattern of trade. The weights employed are 'fixed' base-year weights, determined by the pattern of trade in 1954. Thus the weighting employed for the indices in 1955 and later is given by the pattern of goods traded in 1954. The index itself is derived by calculating the geometric mean of the products of the unit values and their respective weights. In other words, the resultant indices measure the change from period to period in the value of a fixed selection of commodities, regardless of the fact that the composition of the goods traded in any period differs from that in others. This weighting system is adequate only for as long as the pattern of commodity trade remains constant from year to year. If in any particular year there are marked changes, the use of the weights based upon the 1954 pattern will distort the indices in the current period.

Although the price indices are published monthly, separate indices – based of course on the monthly data – are published for successive quarters and for each year. A selection is given in Table 77 for both imports and exports. The annual indices – those for some years back are also reproduced in Table 77 – are used for measuring long period changes in the terms of trade of the United Kingdom. They have replaced the now discontinued *average value* indices (discussed below) which until 1955 were calculated for this purpose. Note that although the current indices are based upon 1954; when the base was changed, comparative indices were calculated back to 1950 to permit comparisons over a longer period. While it is probable that the price indices may before long be revised on to a more recent base year, *e.g.*, 1958 or 1959, it seems unlikely that the construction of the index will change.

Import and Export Volume Indices

Index numbers which measure the changes in the *volume* of both imports and exports are also prepared by the Board of Trade. They are published quarterly with the overseas trade returns in the *Board of Trade Journal*. The indices are designed

to show the variations in imports and exports after eliminating price variations, *i.e.*, volume changes only. This is done by recalculating the value of the imports (or exports) for any quarter at the average prices of the year 1954. By expressing the corrected value of imports (or exports) as a percentage of the 1954 value an index of *volume* change is derived.

The quarterly figures used in calculating the index are derived from the Trade and Navigation Accounts. As with the import and export *price* indices, adjustment of the contents of various headings is necessary. For those items for which only value and not volume is given in the Accounts, estimates of the probable changes are made by assuming that they move in the same manner as do related items for which both value and volume figures are available. This procedure is adopted to a rather greater extent for the export index since a larger proportion of total exports are given in value terms only. Both volume and value figures are available for all but 3 per cent of imports (by declared value).

These indices have been published for many years, but since in the view of the Board of Trade a change in both the base year and structure of the index should be made at least every five years (owing to the changes in the pattern of overseas trade) the index is not comparable as between different base years, except where the change of base has been accompanied by a revision of indices for the earlier years.

The latest base year is 1954 to which the index was revised from 1950 in December 1955.[1] At the same time the weights were revised on the basis of the average prices ruling in 1954 instead of 1950. As a result of these changes the present indices are rather different from the earlier series and direct comparisons would only be possible after the calculation of both series of indices on a common base. In the current index the weights are based on the relative values of the individual categories of goods comprising the total value of trade in the year 1954.

The volume indices shown in Table 78 are also given for the same sub-headings as the import and export *price* indices, *i.e.*, raw materials, manufactures, etc., which were discussed earlier. Like the monthly price indices discussed earlier, the volume indices are separately calculated for each quarter and for each

[1] B.O.T. Journal, 10th December 1955, and 19th May, 1956.

TABLE 78

VALUE AND VOLUME OF UNITED KINGDOM IMPORTS

Class and Division	Value as declared	Estimated Value at 1954 prices	Index numbers of Volume 1954=100					
	Year 1958	Year 1958	Year 1950	Year 1952	Year 1955	Year 1956	Year 1957	Year 1958
TOTAL CLASS A Food, Beverages and Tobacco ..	£m 1,506	£m 1,580	92	91	107	109	113	119
TOTAL CLASS B Basic Materials	908	964	97	90	105	102	106	94
TOTAL CLASS C Fuels	441	407	65	83	121	115	114	124
TOTAL CLASS D Manufactured Goods [c] ..	909	894	83	100	125	125	130	132
TOTAL ALL CLASSES (including Class E miscellaneous)	3,780	3,862	89	92	111	110	114	114

Source: Board of Trade Journal, 13th February, 1959.

year. They are published with the trade figures for each quarter in the *Board of Trade Journal*. One form of their publication is illustrated in Table 78. As before, the indices reproduced are the annual figures for selected years.

Reference was made earlier to the average value indices which were published until 1955 with each quarter's trade return in the *Journal*. The average value index was a form of price index, since it was derived by dividing the declared value of a given quarter's imports by their value at 1950 prices which were always given in the published table. Thus in Table 78 the first two columns of figures relate to the year 1958 showing (1) the declared value of imports in that year and their (2) re-calculated value had their prices remained at the 1954 level. Thus, if the latter, *i.e.*, base year figure is divided into the current declared value, the result is an index of the average change in the prices of all goods imported. Up to 1955 this index was printed quarterly in the tables of the trade figures. It was then discontinued because the composition of imports changed from quarter to quarter and therefore the index reflected not merely price changes, but also any

changes in the make-up of that period's imports or exports. For comparative purposes the authorities prefer to use the annual price indices derived from the monthly indices which, it will be recalled, are based upon a *fixed* collection of goods and weights. In other words, the monthly (quarterly and annual) price indices are base year weighted Laspeyre type indices. The average value index was in effect a current period weighted index of the Paasche type. It was, in fact, the only index of the Paasche variety prepared officially. This was possible because the figures for trade are available very soon after the actual entry or export of the goods. But for the reasons given above, the average value index has been omitted in the published tables, which now appear as in Table 78, but it is still calculated by the statisticians at the Board of Trade for internal use.

The foregoing comments can be illustrated by reference to the data given in Table 78. The average value index for imports in 1958 is given by dividing the 1958 total at its declared value by its estimated 1954 value; *i.e.*, £3,780m. by £3,862m. to give an index of 98. Table 77 gives an index of import *prices* for 1958 as 99; this is the 'monthly' index using 1954 weights, but the difference between the two indices is not large, nor will it be large as long as the pattern of trade, *i.e.*, its composition from year to year, does not alter. The *volume* index for all imports in 1958 is 114 which is derived by converting the declared 1958 imports of £3,780 million by the price index to give their 1954 value which is £3,862 million (the reverse of what was done to get the price index) and then dividing the estimated 1954 value by the imports of 1954 which were actually valued at £3,379 million.[1] The volume index for 1958 imports is 114 as shown in Table 78. Note that the 1954 total of imports is of course the same whether it is 'declared' or '1954 prices'. In the published tables, similar data are reproduced for exports, as well as imports.

Sterling Area Trade Indices

In 1954 the Board of Trade commenced publication of new indices of prices and volume relating to the trade of the sterling area with the outside world. The object of these indices was to provide some measure of the effects of a change in the terms of trade between the sterling area and the outside world, the more

[1] This figure is not reproduced in the above tables, but is given in a table published with other trade figures in the Journal.

so because the sterling area comprises both manufacturing countries and primary producers. It is common knowledge that a decline in the prices of primary products is accompanied by a decline in the export of manufactured goods to primary producers, while a corresponding increase in the prices of primary products usually generates increased demand for manufactured goods. On the other hand, for the manufacturing country the latter situation may give rise to difficulties in respect of an adverse movement in its terms of trade. The two indices described below are an attempt to answer the question how these opposing influences affect the trade of the sterling area as a whole with the rest of the world. In addition indices are calculated showing the effect of these various influences on the trade of the U.K. and the rest of the sterling area respectively with the rest of the world. The division of these indices is clearly brought out in Tables 79 and 80.

Two indices are produced: indices of import and export *prices* to give the terms of trade between the sterling area and the outside world, and *volume* indices of imports and exports. Like the

TABLE 79

STERLING AREA TRADE WITH NON-STERLING AREA

Indices of Prices and Terms of Trade. 1953 = 100

Period	United Kingdom			Rest of Sterling Area			Total Sterling Area		
	Import Prices	Export Prices	Terms of Trade	Import Prices	Export Prices	Terms of Trade	Import Prices	Export Prices	Terms of Trade
1950	88	84	105	90	102	88	89	93	96
1951	116	101	115	110	156	71	113	129	88
1952	116	103	113	109	108	101	113	106	107
1954	98	99	99	96	101	96	97	100	97
1956	104	105	99	99	95	104	102	100	102
1957	106	108	98	103	99	104	105	104	101
1958									
Jan. – Mar.	99	109	92	101	91	112	100	99	101
Apr. – June	97	107	91	100	88	114	99	97	101

Source: Board of Trade Journal, 29th January, 1959.

1954 price and volume index of the U.K. described above, the indices appear quarterly. The compilation of these particular indices was handicapped by the limitations in the currently published statistics of sterling area countries other than the U.K., referred to below as the rest of the sterling area. When the indices first appeared (*B.O.T. Journal*, 23rd January 1954), they were based upon the records of the trading partners of the

sterling area countries rather than on the trade statistics of the sterling countries themselves. In the following year, however (*B.O.T. Journal*, 15th January 1955), the indices were revised so that the price indices for the rest of the sterling area, *i.e.* apart from the U.K., for which the data were already available, were calculated from the trade statistics of the sterling area countries themselves. The base year, too, was changed from 1950 to 1953. In the case of exports the coverage of the index is just under two-thirds of the sterling area's exports in 1953. In the case of the import price index about two-thirds of the data for the rest of the sterling area are still obtained from price indicators of the exports of these countries' trading partners. This introduces some distortion on account of the difference of timing compared with the import data used for the volume indices for the rest of the sterling area, *i.e.* the goods covered by the price and volume indices respectively for any quarter tend to overlap rather than coincide.

The volume indices are derived from the price indices by dividing value relatives of the current trade of the sterling area with non-sterling countries by the corresponding price indices. Since the weighting is based on the quantity pattern of the sterling area trade with non-sterling countries in 1953, it follows that the 1954 volume index numbers are currently weighted. Next year, therefore, the weighting of these indices will be determined by the pattern of trade this year. Table 80 below shows the form of publication employed for these indices.

TABLE 80
STERLING AREA TRADE WITH NON-STERLING AREA
Indices of Volume of Trade

1953 = 100

Period	Imports			Exports		
	United Kingdom	Rest of Sterling Area	Total Sterling Area	United Kingdom	Rest of Sterling Area	Total Sterling Area
1950	102	91	96	104	99	101
1952	96	116	105	99	99	99
1954	104	110	107	101	96	99
1956	124	142	132	124	126	125
1957	129	155	141	126	132	129
1958 Jan.- March	125	155	139	121	134	127
April - June..	125	150	137	121	125	123

Source: Board of Trade Journal.

The Balance of Payments

The Balance of Payments White Paper contains a summary financial record of the overseas trading activities during the past year. Until 1939 it was based upon the data provided by the Customs Statistical Office *i.e.* the Trade Accounts described above. With the introduction in 1939 of Exchange Control, a new basis for these statistics became available. An importer requiring foreign exchange to pay for goods had to make a detailed application to the authorities for the currency. Similarly, the exporter had to account for the proceeds in foreign currency of any exports. The authorities found themselves in possession of far more detailed and accurate data about the nation's overseas financial transactions than ever before. Since actual payment for goods usually takes place after receipt or despatch of the goods concerned, the Balance of Payments account before 1939 was in the nature of a revenue and expenditure account with debits accrued and credits outstanding. When the Exchange Control data became available the account became in effect a cash account, reflecting the timing of the payments rather than the actual movement of the goods giving rise to payments and receipts. When in October 1950 the new series of half-yearly White Papers on the Balance of Payments was introduced the basis was changed. The current accounts record transactions in goods when a change of ownership takes place. In the case of certain imports such a change of ownership takes place in the country of origin, with exports the change is assumed to be effected on or after the arrival of the goods in the port of destination. The balance of payments accounts will differ from the Trade accounts not only in respect of valuation of imports, which for the latter are c.i.f. and for the former f.o.b., but also in respect of timing. There are also differences in the goods covered by the two sets of accounts, *e.g.* precious stones and gold are excluded in the Trade Accounts but included in the White Paper. The latter document contains a table showing the adjustment between the two sets of accounts, but it is clearly of little help in trying to determine the trend of payments when only the monthly figures of visible trade are available.

The Balance of Payments White Paper presents a summary set of accounts showing the nation's economic and financial transactions with the rest of the world. There are in effect four

TABLE 81
General Balance of Payments of U.K.

£ million

A. CURRENT ACCOUNT	1955	1956	1957	1957 January - June	1957 July - December	1958 January - June (provisional)
DEBITS						
1. Imports (f.o.b.)	3,432	3,462	3,573	1,807	1,766	1,616
2. Shipping	341	412	444	234	210	177
3. Interest, profits and dividends ..	269	259	251	124	127	127
4. Travel	125	129	146	55	91	57
5. Migrants funds, legacies and private (net) gifts	18	18	33	17	16	9
6. Government	241	258	248	131	117	135
Total	4,426	4,538	4,695	2,368	2,327	2,121
CREDITS						
7. Exports and re-exports (f.o.b.) ..	3,076	3,411	3,517	1,782	1,735	1,753
8. Shipping	464	517	554	286	268	259
9. Interest, profits and dividends ..	346	373	361	174	187	167
10. Travel	111	121	129	56	73	61
11. Government (*a*) Defence aid (net)	46	26	21	18	3	3
(*b*) Other.. ..	59	65	84	58	26	32
12. Other (Net)	251	283	301	112	189	180
Total	4,353	4,796	4,967	2,486	2,481	2,455
Current Balance of which .	— 73	+258	+272	+118	+154	+334
visible trade	—356	— 51	— 56	— 25	— 31	+137
invisibles: Government.. ..	—136	—167	—143	— 55	— 88	—100
Other .	+419	+476	+471	+198	+273	+297
B. INVESTMENT AND FINANCING ACCOUNT[1]						
Long-Term Capital (Net)						
13. Inter-Government loans by U.K. (net)	— 5	+ 19	+ 13	—	+ 13	+ 9
14. Inter-Government loans to U.K. (net)	— 48	— 70	— 59	— 18	+ 77	— 23
15. (*a*) Other long-term capital (net)..	—130	—190	—280	—160	—120	—100
Total	—183	—241	—208	—178	— 30	—114
Balance of current and long-term capital transactions	—256	+ 17	+ 64	— 60	+124	+220
15. (*b*) Balancing Item	+100	+ 47	+125	+ 77	+ 48	+ 43
Other Financing Items:						
15. (*c*) Miscellaneous capital (net) ..	+ 60	— 70	+ 10	— 20	+ 30	+ 30
16. Overseas sterling holdings of:						
countries	—127	—155	—151	+ 73	—224	+ 1
non-territorial organisations[2] ..	— 7	+200	— 24	— 9	— 15	— 14
17. U.K. debit balance in E.P.U. ..	+ 1	+ 4	+ 11	— 10	+ 21	— 17
18. U.K. Official holdings of non-dollar currencies	—	— 1	— 22	—	— 22	+ 24
19. Special 'waiver' accounts ..	—	— 37	+ 37	+ 37	—	—
20. Gold and Dollar Reserves ..	+229	— 5	— 50	— 88	+ 38	—287
Balance of Investment and Financing . ..	+ 73	—258	—272	—118	—154	—334

[1] Assets: Increase —/decrease +. Liabilities: Increase +/decrease—.

[2] Of which change in I.M.F. holdings in respect of U.K. Drawings: 1956 + 201.

Source: United Kingdom Balance of Payments 1955 to 1958, Cmnd. 540 (H.M.S.O. October 1958).

main accounts, broken down into various detail. The two most important are the account relating to current trade, which is

coupled with the capital account which shows how any surplus or deficit on current trade during the half-year covered by the White Paper is actually financed. The two accounts are termed the Current Account and the Investment and Financing Account respectively; before 1950 the latter was known by the less complex title of Capital Account! In the White Paper these accounts are set one under the other as depicted in Table 81 on page 397 in a series of eight tables. The first table (see Table 81) gives the general current balance of payments while the next seven give the current balance of payments with particular areas of the world, *e.g.*, Sterling Area, Western Hemisphere, O.E.E.C. countries, etc. It will be seen that these sub-divisions of the world are not mutually exclusive. The two remaining main tables in the White Paper show Gold and Dollar Reserves and Overseas Sterling Holdings.

Apart from the ten main tables in the White Paper there are a number of additional notes and tables giving statistical detail. These supplementary tables include figures showing the settlements with the European Payments Union, special receipts and loans from North America, details of government expenditure, films and business travel, etc.

The analysis and breakdown of the accounts is quite detailed but in most tables a fairly large item described either as a 'balancing item' or 'other' appears which detracts from their usefulness. In general the accounts are rather too complex for the layman and short of detail for the expert. A more detailed description of the sources and methods used in compiling the balance of payments was given in the November 1957 issue of *Economic Trends* which the specialist student should consult.

Board of Trade Wholesale Price Index

Between 1951-5 the Board of Trade was producing two indices of wholesale prices. The first, which was the revised version of its original 1921 index, was finally discontinued in 1955. Table 82 shows the annual averages for the constituent groups of that index for the last seven years of its life. The index incorporated some 258 quotations classified into 200 commodities, *i.e.* some commodities being quoted several times to obtain an average. Fully finished articles were not included but were indirectly represented by weighting those semi-finished articles and raw

materials which entered primarily into manufactured articles. The weighting in the index was effected by including in each of the groups of commodities several quotations for particularly significant commodities. The object was that each commodity should be weighted in proportion to its significance in the overall net value of all manufactured goods produced in the United Kindgom as given by the 1930 Census of Production. Each month's average of prices was compared not with that of the preceding month, but with the average of the same month in the preceding year. This ensured that the changes in prices shown by the relatives each month were between the same goods, since many products, for example fruit, are seasonal. The index for the year was the geometric mean of the twelve monthly indices. In other words, the index was of the chain base variety and this fact enabled considerable variations in the constituent items to be made, whilst ensuring that in the short period at least the changes indicated by the index were between comparable sets of prices.

TABLE 82

WHOLESALE PRICE INDEX NUMBERS (OLD SERIES)

(Average 1930 = 100)

Annual Average	Total All Articles	Intermediate Products	Iron and Steel	Total food and tobacco	Cereals	Meat, fish and eggs	Other food and tobacco
1938	101·4	104·5	139·1	97·3	109·9	85·9	97·5
1949	230·0	260·3	252·9	196·7	196·7	156·1	230·6
1950	262·4	294·5	260·8	221·1	235·3	173·6	251·3
1951	319·5	371·8	292·4	246·9	287·0	179·3	278·6
1952	327·6	364·3	353·5	284·1	313·6	222·1	316·4
1953	327·9	355·9	362·1	307·4	341·3	247·0	334·6
1954	329·9	385·9	366·2	307·5	320·1	237·4	358·9

Source: Board of Trade Journal. This series has been discontinued as from 1955 on.

A detailed account of the construction of this index was given in the supplement to the *Board of Trade Journal* of 24th January 1935. This index was replaced by the new index of wholesale prices which appeared initially in the *Board of Trade Journal* of 19th May, 1951. The replacement of the old index was urgently required, since both the composition of the old index and the weighting system, which was based on 1930, were unsatisfactory and out of date. The only reason for including the above account

of the old index is if long period comparisons of price trends are required. The new index of wholesale prices, which is explained below, does not go back earlier than 1950, although some of the indices have been calculated for the years 1946-50 where the data permit.

The old index was to be a means of answering the question 'what is the average change in the value of money relative to other things'; a reflection of the acceptance of the then current quantity theory of money. The new index is based on an entirely different conception of the functions which the index should perform. The new indices reflected the view that there is no such thing as *the* price level. At best the majority of prices move in the same direction, but always in varying degrees. The indices were to be related to major economic groupings, for example industries, and constructed 'as far as possible so that they may be of direct help to the government, to industry and to economists in studying the effects of price changes'.[1]

The new and current index introduced in 1951 is in fact a number of index numbers which have been classified into three main groups. First, there are price indices of commodities and materials which are important in the production processes of certain industries. These commodity price indices relate to materials such as aluminium, brass and copper among metals. Among the staple fibres there is an index for raw cotton which is supplemented by separate indices for five separate types of raw cotton. The index for raw wool is similarly supplemented by three separate indices for the different varieties. The second group of index numbers are to a certain extent based upon the first group; they are termed indices of basic material prices. Among the first of these indices to be produced were those based on the prices of materials used in the mechanical engineering industry and building and civil engineering respectively. It was intended that these particular indices would be sufficiently reliable to permit price revision clauses to be inserted into contracts for public works, whereby an agreed basis for adjustment of prices would be available to contractor and the authority placing the order. The last of the three groups of indices are designed to reflect the price movements of the total output of certain important industries. For example, there is an index for the china

[1] *Board of Trade Journal*, 19th May, 1951.

and earthenware industry, for iron and steel (tubes) and for tinplate. Examples of the indices taken from each of the three groups are given in Table 83 below.

The base date for all the indices is 1954 = 100. The current series of index numbers is based on some 7,000 price quotations as against the 5,000 used when the index was first introduced in 1955 with a base date of mid-1949. The indices are the arithmetic mean of the percentage changes in the prices that have taken place since the base date.

The prices used in the calculation of these indices are the 'ex-works' prices of the commodities. If it is the practice for the industry to quote the price for the commodity 'delivered', then that quotation is used. The weighting is determined by the information derived from the 1948 Census of Production, although supplementary information (the source of which cannot apparently always be disclosed) has also been utilised to obtain correct weighting. In the case of the commodity price indices (group 3) which are compiled upon the basis of a number of types, *e.g.* the raw cotton or wool index, the weighting is determined by reference to the value of the sales of each constituent commodity in 1954. For the basic materials price index (group 1) the weighting is determined by the value of the relevant materials actually consumed in the appropriate sector of industry in 1954. For example, in the case of the house building materials index, bricks form 12·8 per cent of the total weighting, softwood 0·9 per cent and sand and gravel 8·2 per cent.

The indices of prices of output of broad sectors of industry (group 2) derive their weights from data relating to the sales of the output or product by the corresponding sector of industry in 1954. For example, the price index of the output of the iron and steel industry is based upon the combined prices of the commodities contained in the list of commodity price indices, *i.e.* iron castings, sheets, tinplate and tubes. Details of the appropriate weighting has not been published. It should be noted that the prices and weights of the materials used in this particular index, *i.e.* product of broad sectors of industry (group 2), relate only to the output sold outside the industry and not to that sold between firms within the same industry.

The monthly indices are published in the *Board of Trade Journal*. In the mid-July issue each year there is a detailed review

TABLE 83

BOARD OF TRADE WHOLESALE PRICE INDEX:
NEW SERIES 1954 (Annual Average) = 100

	Annual Averages				Monthly Averages 1958			
	1955	1956	1957	1958	Mar.	June	Sept.	Dec.
1. *Materials Purchased by Broad Sectors of Industry*								
Basic materials and fuel used in manufacturing industry	103·0	106·7	107·4	100·8	100·5	101·1	100·2	101·3
Materials and fuel used in the electrical machinery industry ..	110·2	114·3	114·9	114·5	114·0	114·0	114·5	115·1
House building materials ..	105·2	109·4	112·3	111·9	112·4	112·0	111·3	111·5
2. *Output of Broad Sectors of Industry*								
All manufactured products: total sales ..	103·4	106·7	110·2	111·0	110·8	111·0	110·9	111·6
Iron and steel: total sales ..	104·0	112·1	123·9	127·7	129·3	127·8	127·0	126·7
Paper industries: home sales ..	104·9	109·2	110·3	109·7	109·9	109·6	109·4	109·3
3. (*a*) *Commodities produced in the United Kingdom:*								
Coal ..	111·9	127·1	136·5	140·8	142·9	138·9	138·5	142·4
Soap ..	97·7	107·5	115·5	122·1	120·1	122·2	123·4	125·7
Beer ..	100·5	101·8	104·5	105·1	105·0	105·2	105·1	105·1
3. (*b*) *Commodities wholly or partly imported into the United Kingdom:*								
Cotton, raw ..	95·9	95·1	90·7	77·2	80·2	77·6	75·5	70·7
Wool, Merino only: raw ..	83·2	87·7	98·1	69·9	74·6	75·2	64·7	57·9
Wood pulp imported ..	104·9	109·5	107·7	101·9	103·3	102·1	101·8	99·9

Source: Board of Trade Journals – various. February 1959.

of the movements in the indices during the past eighteen months. In the mid-February issue, complete tables for each group of indices showing the annual averages for the latest year together with comparative figures for earlier years are given. Table 83 is a combination of the information given in these issues; it shows

the annual average for selected groups and commodities as well as the monthly indices for selected months in 1958. In the same issue any major changes or additions to the current indices are discussed. A selection of the more important indices in each of the three groups is published in the Monthly Digest of Statistics.

The Actuaries' Investment Index

The most widely recognised index of share and security prices is that prepared monthly by the Institute of Actuaries. The index is prepared by the Institute for circulation to subscribers only but a summary of the most interesting groups and changes each month is usually published in the leading financial journals. As will be seen from the summary table below, separate indices are prepared for a variety of securities and shares. The purpose of the index is to indicate the long-term trend of security prices. Unfortunately for long period comparisons, the index is revised at regular intervals. The last change was made in 1958 when the base date became 31st December 1957. Before this change, the base dates had been end-December 1928, 1938 and 1950. The index is actually a group of indices; a separate index being prepared for particular groups of securities or shares. These are classified into fixed interest stocks and ordinary shares, and for each of these main groups there are a number of indices. Each index is based upon the price movements of several stocks, the smallest group among ordinary shares being the Stores index with only two quotations in contrast with twenty shares comprising the Miscellaneous index.

Government securities are represented by a single security, $2\frac{1}{2}\%$ Consols. In view of the closely integrated structure of the gilt-edged rates this single security, which enjoys a good market, is probably an adequate indicator of price movements of the gilt-edged market. In the case of the group headed 'Home Corporations', four leading corporation stocks, three of which are irredeemable and the other unlikely to be redeemed, are used. The remaining securities and shares indices are classified under three main heads: debenture, preference and ordinary stock. The first of these main groups is subdivided into investment trusts, breweries, and miscellaneous. In the case of debentures as well as separate indices for Brewing and Miscellaneous debentures a combined index has been computed for these two groups.

The prominence of breweries in this group is explained by the fact that their freehold assets, *e.g.* tied houses, constitute an excellent security for loan capital. Preference shares are classified into two main groups, relating to investment trusts and industrials. Ordinary shares are now broken down under three main heads:

(*a*) *Financial*, which include banks, insurance and property company shares;

(*b*) *Equity Stock* of the capital goods industries; and lastly

(*c*) *Equities* issued by companies producing consumption goods.

Separate indices are computed for the chemical, oil and shipping industries as well as miscellaneous. Finally, there is an industrial index covering all equities contained in all the groups. Examples are illustrated in Table 84.

The selection of securities and shares incorporated in the various indices is restricted to those companies the shares of which are quoted in the London Stock Exchange Official List.

TABLE 84

THE ACTUARIES' INVESTMENT INDEX

Securities and Shares	No. of Securities	Price Index (31st Dec. 1957 = 100)					
		7 Dec. 1956	31 Dec. 1957	30 Sept. 1958	28 Oct. 1958	25 Nov. 1958	30 Dec. 1958
BRITISH GOVERNMENT:							
$2\frac{1}{2}$ per cent Consols ..	1	110·8	100	111·8	112·7	110·3	111·8
HOME CORPORATIONS ..	4	112·5	100	108·8	110·8	108·7	108·2
DEBENTURE:							
Breweries	5	112·8	100	108·6	109·3	110·7	111·1
PREFERENCE:							
Investment Trusts ..	10	110·0	100	106·3	107·6	108·3	107·6
Industrials ..	20	102·9	100	102·5	102·7	102·4	102·3
ORDINARY SHARES:							
Industrials – (Capital Goods):							
Building Materials ..	10	106·3	100	134·4	141·1	147·7	159·0
Engineering ..	10	110·3	100	117·6	121·3	123·5	132·0
Total Capital Goods ..	56	110·7	100	117·3	120·7	121·2	131·0
Industrials–(Consumption) Goods:							
Chemists	7	84·4	100	129·3	137·6	137·9	146·6
Food	10	101·5	100	132·3	134·4	138·5	159·2
Total (Consumption Goods	86	98·3	100	123·0	128·9	131·5	141·9
Industrials (All classes combined)	182	103·5	100	123·0	128·9	131·5	141·9

Furthermore, this quotation must have been confirmed by frequent market dealings. Secondly, the share must be issued by a company registered and operating at least in part in Great Britain. In the case of debenture stocks, the nominal issue on the 29th December 1950, must have been at least £500,000; the interest payments thereon must be reasonably well covered by earnings, and the stock should be either perpetual or unlikely to be redeemed. In the case of preference shares, a stock is included in the index only if at 31st December 1956 at least £1,000,000 was in issue in the case of an Investment Trust and twice that figure if issued by an industrial company. Furthermore, it must be irredeemable and carry no participating rights nor carry more than 25% of the total voting strength. The dividend payable on these securities must be cumulative, not tax free, and as with debenture interest must be well covered by current earnings. The ordinary share indices are based upon the prices of those shares which represent the equity of the business. In other words, the shares may actually be deferred or deferred ordinary. Only such equities are included the market capitalisation of which on the 31st December 1956 exceed £2,000,000. Note the valuation is determined by capitalisation, *i.e.* market prices, and not par value. In the case of tobacco and chain store shares the minimum capitalisation is raised to £10,000,000 to render these groups more homogeneous. Finally, the dividend on the shares must have been paid in each of the preceding five years.

It will be apparent from this account of the selected securities that the Actuaries Index is particularly well suited to the institutional investor who is interested, apart from trustee stocks, in high-class preference shares and blue-chip equities with a high degree of marketability.

The computation of the index is straightforward. The prices used are the middle market prices given in the official list on the last Tuesday of each month, while the dividends used in calculating yields are the total dividends paid during the twelve months preceding the date of the calculation. The individual industry index, for example engineering under Industrials (Capital Goods) is the unweighted geometric mean of the price relatives of the securities in that group. In the case of the group index, *e.g.* total Capital Goods, the index is again the geometric mean of the index numbers of the constituent groups. But in the

case of the ordinary share indices, weighting is effected by the market capitalisation of the shares included. Thus, the Engineering index is weighted four times as heavily as shipbuilding in the group index for Capital Goods since the market value of the 10 engineering shares was four times as great at the base date as the market value of 6 shipbuilding companies' shares. The same method of weighting is used for the 'All-Classes' index.

In addition to the price indices the average percentage yields on all classes of shares corresponding to the price index are also given. These are calculated on the basis of the dividend paid in the preceding twelve months.

The index is kept fully up to date and its representativeness ensured by an annual revision of the shares constituting the various indices. Both preference and ordinary shares may be removed if their prices fall below a fixed limit (unspecified) in relation to the index of the group. The revision also permits the inclusion of new securities in the case of ordinary shares which become available as the result of new issues. Similarly, securities may be removed in the event that the company comes under the control of another concern or that it has not conformed to any of the requirements specified above.

Conclusions

The sheer volume of economic statistics published in official returns is impressive at best; at worst it is utterly confusing. The student reader will appreciate that he cannot know them all; in fact a real understanding of any of these data is only acquired after working with them. For example, if one is writing a detailed report on recent trends in British exports, by the end of it, the writer should know his way about the export trade statistics. What is really important is to know what data are available on particular aspects of the economy, then to appreciate that each series has its shortcomings and to look at the footnotes and explanatory notes which usually accompany the tables. Only in this way will mistakes in extraction and consequent errors in interpretation be avoided. The reader who is still wondering what is done with all these and other statistics relating to the economy should read Carter and Roy's book *British Economic Statistics* which although now a little dated still contains an excellent exposition of the problems of utilising some of

these data. More recently a pamphlet published by P.E.P., *Statistics for Government*, No.406, explains how these statistics are used in government economic planning. It should be supplemented by a careful reading of an article 'Recent Developments in Official Economic Statistics' in the May 1957 issue of Economic Trends.

CHAPTER XVIII

TIME SERIES

Introduction

Numerical data, which have been recorded at intervals of time, form what is generally described as a time series. Thus the annual sales of a shop, the quarterly output of coal or the monthly total of passengers carried by a bus company; all these are time series. Undoubtedly the most popular form of presenting such data is in the form of a graph as was shown in Chapter V. Graphs, unfortunately, have only a limited value in statistics; they enable data to be presented in simple and easily intelligible form. They do not, however, add anything to our knowledge of the data and are of no value for analysis, except in so far as a graph does sometimes help to bring out the inter-relationship between two or more time series. For the economist and business man a study of past events is an aid to making judgments concerning the future. Statistical techniques have been evolved which enable time series to be analysed in such a way that the influences which have determined the form of that series may be ascertained.

If the regularity of occurrence of any feature over a sufficiently long period could be clearly established, prediction of probable future variations would become possible within limits. Thus a decline in constructional activity is sometimes regarded as heralding the early stages of a recession. If this assumption can be statistically tested and verified in the light of past experience, then the authorities responsible for economic policy possess a useful piece of knowledge to aid them in their contra-cyclical policy. In practice, the economist and statistician are the first to admit that analysis of trade conditions is extremely complicated since the economic life of a nation is subject to so many complex forces and influences, anyone of which it is impossible to isolate. It cannot be too strongly emphasised, then, that the elementary techniques described in the following text appear deceptively simple. This field of study remains the undisputed preserve of

the experts; and quite recently the validity of some of the generally-accepted and well-established methods of analysis has been questioned.

Types of fluctuation

Most series of economic data may be regarded as composed of four constituent elements which are set out below. In passing it should be noted that not all series combine all four elements, *e.g.* not all trades are seasonal although many are. Note too that not all time series concern economic data. A college may maintain records of examination marks gained by students over their period of study; such a time series would hardly show the same type of fluctuation as that expected from a series based on the quarterly output of motor cars or annual production of electricity. In other words, one applies the methods of time series analysis only to such data as justify their employment, *i.e.* where some useful lessons may be gained for the future. Most economic series covering long periods may be analysed into the following constituent parts:[1]

Types of Fluctuation

(1) The secular trend – or simply the 'trend'.
(2) Cyclical changes.
(3) Seasonal variations.
(4) Irregular or spasmodic fluctuations.

The *secular trend* is the course which the data have followed over a considerable period. Despite temporary deviations from the course, *i.e.* both large and small fluctuations, there is a clearly-marked tendency in a given direction. For example, Table 85 (p. 413) reveals the trend in the amount of income tax collected in the period 1950-58. Although quarterly fluctuations in the series are quite prominent, as can be seen in Figure 21, the *trend* of tax receipts is upward with the passage of time. If a trend can be determined, then the rate of change or progress can be ascertained, and tentative estimates concerning the future made accordingly. The period covered in this example is rather short to justify the term 'trend' which should be restricted to a

[1] Modern statistical theory rejects this highly simplified classification of economic fluctuations. But there is no simpler way of explaining the basic structure of such movements, provided it is remembered that they are all closely interwoven, and their complete separation is virtually impossible Modern analysis emphasises the seasonal movements about relatively short period trends.

definite continuous movement which has been observed over several decades. There is a tendency nowadays, however, to employ the term to indicate the main course followed at the time by the series.

Cyclical fluctuations are far more complex, since their causation differs from period to period. In practice, they are the most difficult of all to anticipate for purposes of effective economic offsetting action. The term 'trade-cycle' would imply a systematic regularity in its appearance, but economists have established a variety of 'cycles', with durations of 3, 5, 7, 9, or 11 years. It is only in recent years that economic opinion has crystallised on the *basic* features of the trade cycle – but despite almost monumental research, particularly in the United States, all that has been proven is that there is no such phenomenon as *the* Trade Cycle, every one is different.

Seasonal variations, however, are somewhat simpler to deal with. It is common knowledge that many industries are more active at certain periods of the year than at others, *e.g.*, the dress and fashion trade anticipates the Spring demand, the toy manufacturers the Christmas season, the motor car industry the Easter and summer holidays, while the building and constructional trades are slack in the winter months. Similarly, indices of grain imports rise during the late autumn due to the American and Canadian shipments of harvested crops, while the demand for agricultural labour is at its peak in harvest periods.

If a definite periodicity for any occurrence can be established, and in particular, the extent of the average fluctuation at that time can be determined, the change in conditions can be anticipated to some degree, and provision made to offset any potential disturbance it might otherwise cause. For example, the Treasury and the Bank of England initiate special offsetting measures to counteract the effect on the clearing banks of such large transfers of tax money during the first quarter of the year as are apparent from the data in Table 85 on page 413.

Irregular fluctuations: The economic life of society would be very much simpler if reliable forecasts concerning the future course of business activity were possible. With existing techniques, forecasting economic trends remains little more than intelligent estimating, since extraneous and unexpected factors continue to appear and upset the best-laid calculations. In the

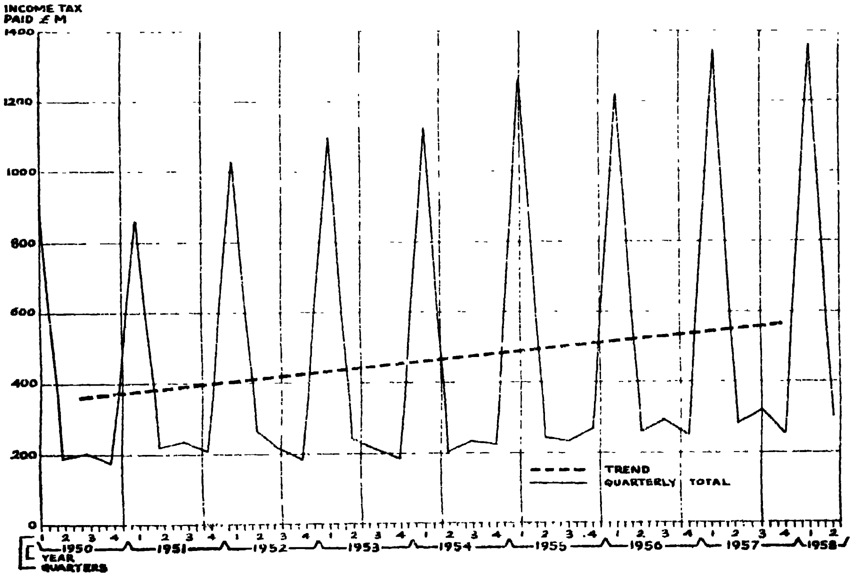
interpretation of any time series, apart from establishing the trend and the extent of the deviations from it, the economist or statistician endeavours to isolate the irregular and unusual fluctuations by determining their cause, *e.g.*, import statistics are made up from Customs records; in the event of a dock strike lasting many weeks, the goods would be landed late and the normal seasonal movement in import statistics would then be replaced by a single swollen figure for the month in which the strike ended. If the seasonal movement were to be computed for a long period, including that year, the distorted figure would have to be adjusted, or even omitted.

Table 85 shows the amount of income tax paid to the Inland Revenue each quarter during the years 1950-8 inclusive. The quarterly totals given in Column 3 are plotted on a linear scale graph in Fig. 21, which conveys almost immediately two clear impressions:

(1) The direction of the plotted curve is upwards from left to right, *i.e.*, throughout the period, the amount of tax paid has increased continuously apart from seasonal variations.

Figure 21

INCOME TAX RECEIPTS QUARTERLY 1950-8 WITH TREND



(2) There appears to be a regular rise and fall in the amount of tax paid within the space of each year, and throughout the whole period these movements appear almost identical from one year to another.

At the outset it can be stated that the first observation concerns the *Trend;* the second refers to the distinct *Seasonal* movement characterising this series. As a first approximation, the trend line may be drawn 'free-hand', sketching through the fluctuation so that the minor seasonal fluctuations are ignored; thereby indicating the course of the data in the period. This method unfortunately requires not merely a moderate degree of artistic skill, but a considerable knowledge of the data and relevant statistical techniques. Not all series are as simple and clear-cut as that depicted in Fig. 21, and the trend line for many series when fitted by different people will vary quite a lot.

The Moving Average Method

The more usual method is to obtain a number of plotting points through which the trend line should pass. These points are obtained by the method set out in Table 85, Cols. 4-6. This involves selecting a number of consecutive values and averaging them so that the variations in the individual values are reduced. The number of values utilised for determining the average depends on the periodicity of the fluctuations. The periodicity of these movements is usually measured by the time between the recurrent 'peaks' shown on the graph of the original values. In this example no real problem of selecting the correct period can be said to exist, the data clearly requiring the average of four consecutive values. In contrast, however, for many other series, the determination of this figure may be extremely difficult. Thus, data relating to business activity, *e.g.*, a series giving the number of company liquidations over several decades, would cover a number of business 'cycles'. Probably it would be necessary to experiment with 5, 7, 8, or 9 years moving averages before a suitable trend line could be clearly established, *i.e.*, when the fluctuations of the individual values in the series from the trend are reduced to a minimum.

If the four values, *i.e.*, those for the four quarters of 1950, are averaged, the average would be written between the values of the second and third quarters. This necessitates the technique

TABLE 85

AMOUNT OF INCOME TAX PAID IN EACH QUARTER 1950–58 (£ million)

(1) Year	(2) Quarter	(3) Quarterly Totals	(4) Sum of (3) in Fours	(5) Sum of (4) in Pairs	(6) Centred Trend (5) ÷ 8	(7) Total Fluctuations from Trend (3) — (6)	(8) Seasonal Variations	(9) Residual Fluctuations
1950	March ..	876	—	—	—	—	—	—
	June ..	185	1,424	—	—	—	—	—
	September	194	1,405	2,829	354	— 160	— 216	+ 56
	December	169	1,435	2,840	355	— 186	— 251	+ 65
1951	March ..	857	1,472	2,907	363	+ 494	+ 675	— 181
	June ..	215	1,507	2,979	372	— 157	— 209	+ 52
	September	231	1,669	3,176	397	— 166	— 216	+ 50
	December	204	1,714	3,383	423	— 219	— 251	+ 32
1952	March ..	1,019	1,697	3,411	426	+ 593	+ 675	— 82
	June ..	260	1,675	3,372	422	— 162	— 209	+ 47
	September	214	1,736	3,411	426	— 212	— 216	+ 4
	December	182	1,715	3,451	431	— 249	— 251	+ 2
1953	March ..	1,080	1,710	3,425	428	+ 652	+ 675	— 23
	June ..	239	1,713	3,423	428	— 189	— 209	+ 20
	September	209	1,752	3,465	433	— 224	— 216	— 8
	December	185	1,713	3,465	433	— 248	— 251	+ 3
1954	March ..	1,119	1,735	3,448	431	+ 688	+ 675	+ 13
	June ..	200	1 761	3,496	437	— 237	— 209	— 28
	September	231	1,873	3,634	454	— 223	— 216	— 7
	December	211	1,913	3,786	473	— 262	— 251	— 11
1955	March ..	1,231	1,911	3,824	478	+ 753	+ 675	+ 78
	June ..	240	1,963	3,874	484	— 244	— 209	— 35
	September	229	1,945	3,908	489	— 260	— 216	— 44
	December	263	1,965	3,910	489	— 226	— 251	+ 25
1956	March ..	1,213	2,022	3,987	498	+ 715	+ 675	+ 40
	June ..	260	2,010	4,032	504	— 244	— 209	— 35
	September	286	2,132	4,142	518	— 232	— 216	— 16
	December	251	2,153	4,285	536	— 285	— 251	— 34
1957	March ..	1,335	2,193	4,346	543	+ 792	+ 675	+ 117
	June ..	281	2,185	4,378	547	— 266	— 209	— 57
	September	316	2,221	4,406	551	— 235	— 216	— 19
	December	253	2,239	4,460	557	— 304	— 251	— 53
1958	March ..	1,371	—	—	—	—	—	—
	June ..	299	—	—	—	—	—	—

of 'centering' the moving average, which can be done by next finding the average of the second, third, fourth and fifth items, and adding this to the average of the first four values. Then if the aggregate of the two averages is halved, the moving average will be 'centered', *i.e.*, the average of the two averages will lie against the third figure of the series, instead of between the second and third. If the period averaged covers an odd number of months or years (almost invariably the latter), the problem of centering does not arise; the average lies against the middle item, *e.g.*, a nine-yearly average would be placed against the fifth value. It should be noted that 'centering' is only really necessary if the calculation of the extent of the seasonal or cyclical fluctuations is required; it is not absolutely necessary for the sole purpose of drawing the 'trend' line, as plotting points can be derived from the 'mid-values'.

A simpler way of working, used in Table 85, is to aggregate the first four values, then the second to fifth values inclusive, yielding successive totals of 1,434 and 1,415 (Col. 4). By adding them, as in Col. 5, and dividing by 8, the moving average is centred (Col. 6). By continuing this process throughout the series, the succession of averages in Col. 6 is obtained.

The series of moving averages plotted on the graph (Fig. 21) yields the dotted line passing through the graph of the original data. This dotted line is described as the trend or line of trend. The process of averaging leaves both ends of the trend line 'short', but this is not important if the period is long.

Derivation of Seasonal Movement

The purpose of analysing time series is not always the determination of the trend. Interest may be centred on the seasonal movement displayed by the series and, in such a case, the determination of the trend is merely a stage in the process of measuring and analysing the seasonal variation. If a regular basic or underlying seasonal movement can be clearly established, forecasting of future movements becomes rather less a matter of guesswork and more a matter of intelligent forecasting. Thus, if grain shipments are known to fall 30 per cent from the August level by the following December, provision for this change and all that it implies for ports, etc., can be made by those concerned. If in the event the fall is sharper still, there is still the consolation that some part of it was correctly forecast. Lastly, before proceeding to the actual statistical techniques, it should be emphasised that the conditions in all the periods of time for which data are available were probably different. The decision as to what constitutes *normal* and the extent of the 'abnormal' fluctuation may be quite arbitrary.

Where, as with the data given, the seasonal movements are large, it is necessary to estimate the 'normal' seasonal movement at each quarter, and show up the 'residual' variation, *i.e.*, the abnormal or unexpected movements. The first stage in this process is completed by averaging the actual movements or deviations from the trend (given in Table 85, Col. 7) for each quarter over the whole period. Since only two quarterly figures of fluctuations are available for 1950, they have been excluded from

the following table showing the method of calculating the average seasonal fluctuations.

Before the average seasonal fluctuation can be calculated and graphed, a further adjustment is usually necessary. It may be remembered that the sum of the individual deviations of a series of values from the mean of that series was equal to zero. Since the line of trend is also an average, the sum of the fluctuations from that line should also equal zero. If they are totalled, as in

TABLE 86

QUARTERLY FLUCTUATIONS

Year	Quarters			
	First	Second	Third	Fourth
1951	+ 494	— 157	— 166	— 219
1952	+ 593	— 162	— 212	— 249
1953	+ 652	— 189	— 224	— 248
1954	+ 688	— 237	— 223	— 262
1955	+ 753	— 244	— 260	— 226
1956	+ 715	— 244	— 232	— 285
1957	+ 792	— 266	— 235	— 304
Total	+ 4,687	— 1,499	— 1,552	— 1,793
Less Bias* ..	+ 39	+ 39	+ 39	+ 39
	+ 4,726	— 1,460	— 1,513	— 1,754
Seasonal Variation..	+ 675	— 209	— 216	— 250

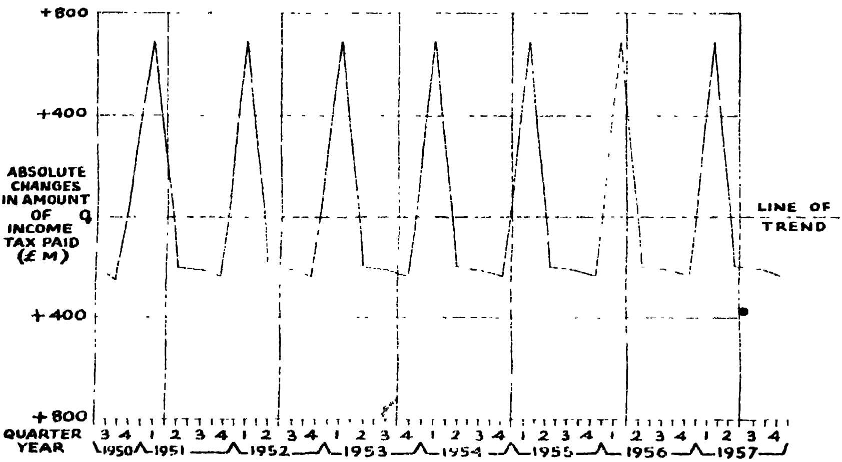
* Note that the error in total is — 157, therefore the correction to be applied is + 157 or + 39 to each quarter.

Table 86 (+ 4687 — 1499 — 1552 — 1793 = 4687 — 4844) there is a difference of — 157 between the sums of the positive and negative movements. This difference, described as a 'bias', frequently arises since the effects of large but irregular fluctuations cannot be entirely eliminated by the moving average method. Since the source of the 'bias' is unknown, it is offset by averaging it over the four quarters, and in this case adding it to each quarter's total. The net sum of each quarter's fluctuations from the trend line is then divided by the number of years covered; the quotients represent the *average* seasonal variation which may reasonably be expected in each respective quarter provided the conditions prevailing during the period covered by

the data do not change unduly. The residual movements have then to be explained by the variety of extraneous influences which have affected the data. The amplitude of the seasonal movement is indicated graphically in Fig. 22. The deviations from the trend line, represented in this diagram by the horizontal axis marked as 0, are measured absolutely along the ordinate.

Figure 22

AVERAGE QUARTERLY FLUCTUATIONS IN INCOME TAX RECEIPTS 1950-7.



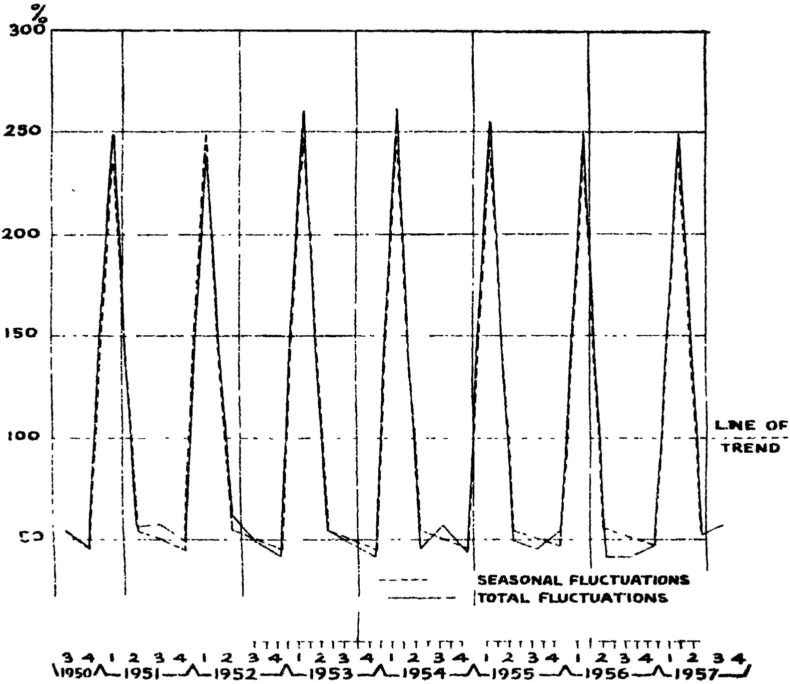
Indices of Seasonal Movements

The movements about the trend line may also be graphed relatively to the trend, *i.e.*, as percentage changes from the trend values at various points of time. This is done in Fig. 23. Such a graph reveals whether the amplitude of the fluctuations is declining relatively to the average values, *i.e.*, the trend. In the foregoing example, although the trend reflects the increasing amount of income tax paid in successive years the seasonal movement has remained relatively stable, *i.e.*, the proportions of the total amount of tax paid made in each quarter have not varied noticeably over the eight years under scrutiny. The total fluctuations (Col. 7, Table 85), *i.e.*, seasonal plus residuals (Cols. 8 and 9), are superimposed on the same graph of seasonal variations (Col. 8).

To derive this graph each quarterly total of income tax paid is

Fig. 23

QUARTERLY FLUCTUATIONS RELATIVE TO TREND: DATA AS IN TABLE 87.



divided by the trend value for that quarter and the result expressed as a percentage of the trend value. Thus the values expressed for September quarter 1950 are $\frac{194}{354} \times 100$ (*i.e.*, trend value divided into the actual quarter's value) equals 54·8; the same calculation being performed for each quarter's figures. The use of logarithms, which are set out in Cols. 4 to 6 of Table 87 simplifies the calculation. The difference between the logarithms represents the index derived by dividing the trend values into the recorded quarterly revenue figures, *i.e.*, Cols 3 $\div$ 6 Table 85.

By averaging the figures for each quarter in the years under review, *i.e.*, first quarter: 236·1, 239·1, 252·3, 259·5, 257·5, 243·5, 245·9, average = 249 to the nearest unit after correction for bias, an index of the seasonal movement may be derived. The sums of the indices for the four quarters should be 400; just as the deviations of the observed values should have equalled zero

TABLE 87

CALCULATION OF SEASONAL INDICES

(1) Year	(2) Quarter	(3) Quar- terly Totals	(4) Trend Values	(5) Logs of Original Values (Col. 3)	(6) Logs of Trend Values (Col. 4)	(7) Differ- ences between (Cols. 5 and 6)	(8) Anti- Logs of Quar- terly Indices × 100
		£m	£m				
1950	March ..	876	—	—	—	—	—
	June ..	185	—	—	—	—	—
	September	194	354	2·2878	2·5490	$\bar{1}$·7388	54·8
	December	169	355	2·2279	2·5502	$\bar{1}$·6777	47·6
1951	March ..	857	363	2·9330	2·5599	0·3731	236·1
	June ..	215	372	2·3324	2·5705	$\bar{1}$·7619	57·8
	September	231	397	2·3636	2·5988	$\bar{1}$·7648	58·2
	December	204	423	2·3096	2·6263	$\bar{1}$·6833	48·2
1952	March ..	1,019	426	3·0080	2·6294	0·3786	239·1
	June ..	260	422	2·4150	2·6253	$\bar{1}$·7897	61·6
	September	214	426	2·3304	2·6294	$\bar{1}$·7010	50·2
	December	182	431	2·2601	2·6345	$\bar{1}$·6256	42·2
1953	March ..	1,080	428	3·0334	2·6314	0·4020	252·3
	June ..	239	428	2·3784	2·6314	$\bar{1}$·7470	55·9
	September	209	433	2·3201	2·6365	$\bar{1}$·6836	48·3
	December	185	433	2·2672	2·6365	$\bar{1}$·6307	42·7
1954	March ..	1,119	431	3·0487	2·6345	0·4142	259·5
	June ..	200	437	2·3010	2·6405	$\bar{1}$·6605	45·7
	September	231	454	2·3636	2·6571	$\bar{1}$·7065	50·9
	December	211	473	2·3243	2·6749	$\bar{1}$·6494	44·6
1955	March ..	1,231	478	3·0902	2·6794	0·4108	257·5
	June ..	240	484	2·3802	2·6848	$\bar{1}$·6954	49·6
	September	229	489	2·3598	2·6893	$\bar{1}$·6705	46·8
	December	263	489	2·4200	2·6893	$\bar{1}$·7307	53·8
1956	March ..	1,213	498	3·0838	2·6972	0·3866	243·5
	June ..	260	504	2·4150	2·7024	$\bar{1}$·7126	51·6
	September	286	518	2·4564	2·7143	$\bar{1}$·7421	55·2
	December	251	536	2·3997	2·7292	$\bar{1}$·6705	46·8
1957	March ..	1,335	543	3·1255	2·7348	0·3907	245·9
	June ..	281	547	2·4487	2·7380	$\bar{1}$·7107	51·4
	September	316	551	2·4997	2·7412	$\bar{1}$·7585	57·4
	December	253	557	2·4031	2·7459	$\bar{1}$·6572	45·4
1958	March ..	1,371	—	—	—	—	—
	June ..	299	—	—	—	—	—

(Table 88). In this case, too, any correction is apportioned between the indices for each of the four quarters. The indices are then rounded.

By applying the indices so obtained to the quarterly totals of another series subject to the same quarterly movement, a new series is derived, *i.e.*, one 'adjusted for seasonal variations'. In brief, the new series then contains the values which would have

TABLE 88

CALCULATION OF SEASONAL INDICES

Year	Quarters			
	First	Second	Third	Fourth
1951	236·1	57·8	58·2	48·2
1952 . ..	239·1	61·6	50·2	42·2
1953	252·3	55·9	48·3	42·7
1954	259·5	45·7	50·9	44·6
1955	257·5	49·6	46·8	53·8
1956	243·5	51·6	55·2	46·8
1957	245·9	51·4	57·4	45·4
Total	1,733·9	373·6	367·0	323·7
Less Bias	0·5	0·5	0·5	0·5
	1,733·4	373·1	366·5	323·2
Seasonal Variation..	249	53	52	46

appeared if the seasonal movement did not exist. This is one of the simplest means employed to 'refine' data subject to cyclical and seasonal variations. If indices can be computed for all variations and then applied to the relevant time series, the analysis is rendered more effective.

Summary

The moving average method of deriving the trend of values and the seasonal fluctuations over a period is the main elementary method available for this work. It can be applied to series covering several decades provided the data can still be treated as a series, the moving average in such a case being used to eliminate the so-called Trade Cycle. When the cycle is regular both in its periodicity and in the amplitude of fluctuations the moving average method will completely eliminate the fluctuation and yield a straight trend line. Skill and knowledge in the selection of the period used for averaging is necessary where the cycle is not so clearly defined, *e.g.*, is the cycle to which the data under scrutiny are subject one of 3, 5, 7, 9, or 11 years; how many cycles are superimposed on one another?

The moving average method, however, has certain disadvantages:

(*a*) The trend line cannot be accurately established for the period covering the whole series. As already stated, the trend line falls 'short' at both ends, and if the cycle is one of, say nine or eleven years, this may constitute a marked gap where the data cover only two or three cycles.

(*b*) The difficulty already explained of establishing a definite periodicity in the fluctuations. Different views on this will result in differing trend lines. Unless, therefore, the seasonal or cyclical movement is definite and clear-cut, the moving average method of deriving the trend may be rather unsatisfactory.

(*c*) Since the 'trend' values are arithmetic averages, any extreme individual variation affects them unduly. If the seasonal variations vary considerably in extent from year to year, the trend may appear as a series of humps rather than a smoothed line, *i.e.*, it is not possible to entirely eliminate the seasonal variation in the series.

Conclusions

Time series analysis requires almost more knowledge of the data and relevant information about their background than it does of statistical techniques. Whereas the data in other fields may be controlled so as to increase their representativeness, economic data are so changeable in their nature that it is usually impossible to sort out the separate effects of the various influences. Attempts to isolate cyclical, seasonal and irregular, or random movements, are made in the hope that some underlying pattern of change over time may be brought out. It is noteworthy that analysis of cycles up to 1914 does reveal what appear to be regular waves of alternating boom and slump. Some of the older trade cycle theories, such as the sun-spot and harvest variations, which are themselves natural phenomena, recognised that economic activity seems to follow a rhythmic pattern. But in present times, when natural economic forces are repressed and government action is so widespread in its effect, the fluctuations of the economy become increasingly unpredictable.

The cyclical pattern of activity revealed over the major part of the last century, on which some of the above statistical ideas are based, gave an impetus to the use of statistical analysis for forecasting trends and cycles. Despite the use of both simple and

complex techniques, ranging from the simple averaging of seasonal movements to multiple correlation of related series, *e.g.*, coal output, rail transport, and steel production, it remains a regrettable fact that 'scientific forecasting' has not added any laurels to statistical science. The ideal situation would be where 'indicators' were available which could be used for forecasting the direction and amplitude of future changes. Both economists and statisticians, by examining and analysing business data over many decades have tried to determine which indicators are the most reliable and consistent. Whether the results justify the volume of work entailed remains to be seen.

Fitting the Trend by Algebraic Methods

A more exact method of defining the trend is to estimate it by algebraic methods.

In Chapter XII on Correlation, it was stated that linear distributions could be summarised by a simple equation of the first degree: $y = a + bx$. This tool is extensively employed for fitting the trend to time series. As in linear correlation where the relationship is represented by a straight line, the equation is of the first degree. Many time series when plotted (as well as variables in correlation analyses) are curvilinear, the curve then being described by more complex equations. The principle, however, remains the same.

To illustrate the technique a simple example is given below. These data were selected because they conform to a linear pattern, but the results of using this method are not very much

TABLE 39

PASSENGER MILES FLOWN IN SCHEDULED FLIGHTS FROM U.K. AIRPORTS 1945–57
(Million miles)

Year	Passenger-miles flown	Year	Passenger-miles flown
1945	301·9	1952	1,242·5
1946	362·8	1953	1,434·2
1947	441·1	1954	1,515·4
1948	554·5	1955	1,801·4
1949	614·7	1956	2,102·3
1950	794·0	1957	2,431·1
1951	1,065·0		

(*Source: Statistical Abstract No.* 95)

better than the Moving Average method. The trend line, however, does cover the whole period and where the relationship is curvilinear, the error which arises in the Moving Average method is avoided. Briefly, when the curve is convex upwards, the trend values derived by the Moving Average method are low; when the curve is concave downwards, they tend to be exaggerated.

The data in Table 89 above illustrates the growth in the air passenger traffic of the United Kingdom between 1945 and 1957.

For quite a considerable body of phenomena, even in the economic field, data may reflect a continuous and fairly consistent rate of change or inter-relationship which may occasionally be described by a mathematical equation. Too much should not be read into this. No economic data will ever conform exactly to some scientific law expressed mathematically. Some data, however, do conform approximately to this concept and for certain analytical work, *e.g.*, statistical testing of trade cycle theories, it is useful to be able to define the trend in such a form instead of using the foregoing moving average method which merely

TABLE 90

CALCULATION FOR FITTING TREND LINE TO TIME SERIES

(Data from Table 89)

(1) Year	(2) Passenger-miles flown (to nearest million) y	(3) Years in sequence x	(4) Col. 2 × Col. 3 xy	(5) x^2	(6) y'
1945 ..	302	1	302	1	76·9
1946 ..	363	2	726	4	252·1
1947 ..	441	3	1,323	9	427·2
1948 ..	555	4	2,220	16	602·4
1949 ..	615	5	3,075	25	777·5
1950 ..	794	6	4,764	36	952·6
1951 ..	1,065	7	7,455	49	1,127·8
1952 ..	1,243	8	9,944	64	1,302·9
1953 ..	1,434	9	12,906	81	1,478·1
1954 ..	1,515	10	15,150	100	1,653·2
1955 ..	1,801	11	19,811	121	1,828·3
1956 ..	2,102	12	25,224	144	2,003·5
1957 ..	2,431	13	31,603	169	2,178·6
Totals ..	14,661	91	134,503	819	

describes instead of analysing the data. Especial care is required, and this frequently constitutes the stumbling block to the use of this technique, to ensure that the basic conditions are unchanged throughout the period to which the data relate. Thus in the example depicting the growth of air passenger traffic the period 1945-1957 is selected rather than say 1937–1949 since the war years can hardly be linked with the post-war period to judge the degree of expansion and the probable future trend.

Method of Least Squares

In order to 'fit' a curve to the data given in the above Table the values of the terms a and b in the equation $y = a + bx$ must be computed. The method employed is detailed in the calculations below and is known as fitting a curve by the 'method of least squares'. The best fit is obtained when the sum of the squares of the deviations from the trend line is at a minimum; hence the name of the method. There is only one line which meets this condition.

The method employed in Table 90 above can be set out in successive stages.

(1) The original data, the years within the period under review, and the annual totals are set down in columns 1 and 2.

(2) Col. (3) headed 'x' represents the 'plot' along the abscissa (horizontal axis) and is written as a consecutive series 1 to 13 inclusive. 'x' may be referred to as the independent variable.

(3) 'y' in the equation is the 'dependent variable', i.e., the annual miles flown. In Col. (4) headed 'xy' the products of Cols. 2 and 3 are given, x in this case being not the year but its position in the series 1 to 13.

(4) In Col. 5 the values in Col 3, *i.e.*, the years in sequence 1–13 are squared.

(5) The values in Cols. 2, 3, 4 and 5 are next totalled, and the aggregates are then substituted for the symbols in the following equations which are employed to derive the trend line:

(i) $\Sigma y = na + b\Sigma x$

(ii) $\Sigma xy = a\Sigma x + b\Sigma x^2$

Thus the equations become:

(i) $14{,}661 = 13a + 91b$

(ii) $134{,}503 = 91a + 819b$.

(6) The solution of these simultaneous equations proceeds as follows: either a or b must be eliminated to derive the value of the other symbol. By multiplying the first equation by 7 the following results are obtained:

(i) $102{,}627 = 91a + 637b$

(ii) $134{,}503 = 91a + 819b$.

The difference between the two equations:

$31{,}876 = 182b$.

$\therefore b = \frac{31876}{182} = 175{\cdot}14$.

(7) Substituting 175·14 for b in the original equation:

$$\begin{aligned}
14{,}661 &= 13a + 91b \\
14{,}661 &= 13a + 91\ (175{\cdot}14) \\
14{,}661 &= 13a + 15{,}937{\cdot}74 \\
14{,}661 - 15{,}937{\cdot}74 &= 13a \\
-\ 1276{\cdot}74 &= 13a \\
-\ 98{\cdot}21 &= a
\end{aligned}$$

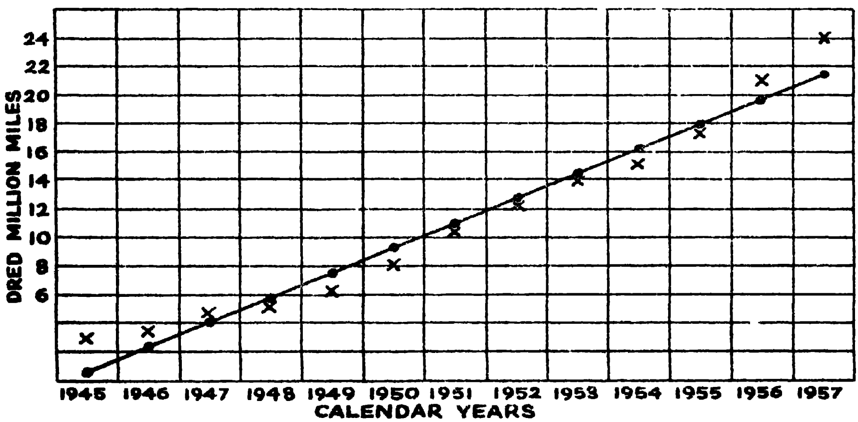
(8) The basic equation which yields the straight line curve is calculated by substituting for a and b the values —98·21 and 175·14 respectively. Thus for 1945, the value of y' is derived as follows

$$y' = -98{\cdot}21 + 175{\cdot}14 = 76{\cdot}93$$

and for 1954: $y' = -98{\cdot}21 + 175{\cdot}14\ (10) = 1653{\cdot}19$

Fig. 24

LINE OF TREND FITTED BY METHOD OF LEAST SQUARES: DATA FROM TABLE 89



The estimated values (y') for each successive year are set down in Col. 6 of Table 90 and may be compared with the original values given in Col. 2. Fig. 23 shows the original values plotted as crosses on the graph; the straight line passing through the positions given by the values in Col. 6.

For future reference, however, it will be necessary only to write; 'Trend of Passenger-miles flown from United Kingdom Airports in million miles 1945–57 $y = -98{\cdot}21 + 175{\cdot}14x$. Year of Origin 1944', a statement which not only describes the character of the data, so that the approximate mileage flown in any intermediate year can be calculated, but also defines precisely the trend.

It will be realised that only two values for y' are required for the drawing of a straight line. The values for y' for the whole range of x values are calculated only so that the annual fluctuations from the trend may be derived, as well as enabling the actual values for each year to be compared with the 'trend'.

The student may wonder why in the note concluding Chapter 12 explaining the calculation of the equation of a regression line, an apparently different method was used from that above to obtain the line of trend. Actually the same result can be obtained by either method; but for calculating an estimated line of trend it is simpler to use the above rough and ready method.

CHAPTER XIX

STATISTICS IN BUSINESS AND INDUSTRY

The efficiency of any industrial or commercial undertaking is ultimately dependent upon effective management. Furthermore, of all the bottlenecks which limit the expansion of the business organisation (and the government department), in the long run management is the most serious. When concerns were small, the director had all the facts at his fingertips and made his decisions in the light of those facts. Today, the size of the public corporation and industrial undertaking is such that management can only follow developments, both internal and external, which affect their production and markets, by the creation of a statistical unit. The function of such a unit is to feed management with relevant data on the basis of which decisions are made, not blindly, but in the light of known facts.

This work may not always be entrusted to a statistical department proper; sometimes a small unit is attached to the accounts department, or more often to the sales director's office. Occasionally an individual with some statistical and economic training is appointed as assistant to the managing director, so that the latter may be kept abreast of developments which might otherwise escape his notice, or provide him with data relating to any problem which he seeks to elucidate. But whatever the form of the statistical 'unit', its functions are basically the same. It provides data which may be broken down into two main streams. Internal data based upon the records maintained by the sales, accounting, production and personnel manager's departments. External data which affect the concern; *e.g.* statistics relating to the trend of trade and of prices. In this connection a knowledge of sources and the reliability of the published data is especially important. An interesting example of the information about an industry and its market which can be assembled almost entirely from official statistics is provided by the Statistical Review prepared by the Furniture Development Council as a supplement to its annual report.

Internal Statistics

In practice, the type of statistical work performed will largely depend on the nature of the business. If its sales are strongly seasonal, the statistician may be able to estimate these variations accurately enough to effect economies on the production side. Similarly, intensive study of the data may reveal relationships with the experience of other industries which may occasionally prove helpful as guides in the taking of policy decisions. In other words, the statistician himself decides the scope of his work, and must use his specialist knowledge and skill to ensure that his efforts provide results which will be of value to the organisation. His biggest enemy is likely to be the director who wants 'facts', not the qualified statement which the statistician must invariably make when his estimate is based either on a sample or on assumptions relating to the trend of business. Even at the expense of disagreement, it is better in the long run not to be bullied into stating positive and unreserved conclusions except on the very few occasions when they are justified.

An important part of the business statistician's work entails the analysis of data and information provided mainly by the accounts and sales departments. From the statistical viewpoint, the sales ledger alone conceals a wealth of information. Turnover can be analysed by areas, by size of customers' orders, by type of goods sold, duration of credit granted or extracted. Similarly, data relating to selling costs, advertising, transport and commissions may be broken down. The internal administrative costs of the organisation may be analysed by departments; or various costs may be related to turnover or profits. The value of such analysis is reflected in the fact that the first task undertaken by any firm of business consultants called in to inquire into the efficiency and marketing problems of any business is to extract all the data possible from the firm's internal records.

The need to simplify statistical data in business cannot be sufficiently emphasised. The statistician who lays before a departmental chief or management board a mass of data, even if it is relevant to the topics on the agenda, is useless. For this type of work a good knowledge of tabulation is essential. In dealing with non-statisticians, the design and lay-out of tables and diagrams are probably as important as the content of the tables, and the

lessons to be learnt from them. The statistician's primary function may be to collect these data, but his real function is to ensure that the board receive a concise statement as to what the figures show. It is his function to interpret his data; he is not a clerk employed to prepare tables of figures. Just sufficient figures should be attached to the report to illustrate the main points. He should, however, keep in reserve all the detailed data, condensed in varying degrees, so that he may produce further and more detailed evidence to support his report in the event that anyone should query it. It has been well put that 'his soundest deductions will be dismissed by that frightful crack about lies, damned lies, and statistics. The human tale of a single customer may outweigh the statistical evidence of ten thousands'.[1] In practice, it is better for a statistician or the statistical unit to be independent of departmental responsibilities. Once 'attached' to a department, the statistician often ceases by sheer force of circumstances to be an independent observer, and is all too often used to support his department's case in interdepartmental disputes.

Labour problems constitute at the present time one of industry's biggest headaches. In this field, too, the statistician can be useful to management. Quite apart from providing the personnel manager with regular information concerning labour turnover, which at any time, much less in conditions of full employment, is expensive for the firm concerned, statistics may be assembled to provide data relating to absenteeism and sickness rates. It may be argued that such data are quite easily assembled by the personnel manager without the aid of a statistician. Just recently, however, the writer has heard of two cases, both concerning nationally known firms which pride themselves upon their labour relations. In one case male and female employees were 'lumped together' for the determination of sickness rates, while in the other case the company seemed to be unaware that there was any relationship between sickness and age, or whether married women, who formed a significant proportion of the staff, experienced different rates from their unmarried sisters. Consequently no analysis of absent or sick employees was made and useful information was lost. These may seem very obvious points to the reader. One is indeed sometimes

[1] 'Statistics and the Statistician in Business', H. T. Weeks, *J.R.S.S.*, 1947, Part II.

bewildered by the omissions and errors made when the statistically untrained individual (and it must be admitted, the statistically trained too) handles or assembles any data. Yet it is in these simple data that improvement is so easy and is so often needed.

External Statistics

'Desk research', as it is often called, does not consist solely of collecting data relating to the activities of the firm in the fields of production and of selling. Reference has already been made to the use of national statistics prepared by the government departments, which – as pointed out in the address of Lord Heyworth, Chairman of Unilever Ltd.[1] – are useful in enabling management to formulate policy decisions in the light of relevant information. A company selling a consumer product needs to watch the national income statistics relating to consumer expenditure which appear annually in the Blue Book and quarterly in the *Monthly Digest of Statistics*. Producers of capital goods can derive useful information regarding the state of the market both actual and potential by studying current investment figures such as the rate of factory building, the state of order books in the ship building industry and that of the machine tool trade which appear regularly in the *Board of Trade Journal*. In the field of exports, statistics are more than adequate to meet most needs for guidance as to the changing pattern of demand in leading overseas markets. The Board of Trade is prepared to supply detailed analyses and breakdowns of statistics relating to particular goods and markets for manufacturers and others who require information in greater detail than in its published form.

The firm contemplating the marketing of a new product can obtain information as to the number of retail outlets which carry this type of goods throughout the country and their distribution. It is on the basis of the Census of Distribution that the advertisers of new shop premises in a Yorkshire town were able to advise readers to establish themselves in Britain's 'most undershopped town'. In the same field, fluctuations in the monthly retail trade statistics can be used as a standard against which the firm's own sales performance can be judged in general terms. Where the market for a particular commodity is dominated by a

[1] See quotation on page 7.

mere handful of large firms, a study of changes in Censuses of Production data can reveal the extent to which the numerous small firms share in an expanding market. It is easier for the large firm to increase its sales by pushing the smaller firm out of the market, than by trying to beat its large rival. By no means statistical in character, but nevertheless often full of valuable information are the chairman's remarks at the annual general meetings of the leading public companies. Firms whose prosperity is dependent upon the sale of their particular products to such firms or ancillary trades can often derive some guidance as to their prospects from the declared intentions of their main customers. In brief, there is an adequate supply of economic statistics and even a superfluity thereof in some fields, which if intelligently used can ease the task of management in regulating the affairs of its business.[1]

A major problem confronting the statistician in business is to answer the question, '*What* figures should be produced?' It is all too easy to produce masses of detailed information, most of which will never be used. It is even easier to continue producing figures the need for which has long passed. It is important, however, to ensure that sufficient data are available to answer promptly any request for information on any aspect of the organisation's work. As a basic principle, it should never be forgotten that figures cost money to produce. It is a sheer waste to produce more than are absolutely essential to the efficient conduct of the business. As one industrial statistician has put the problem: 'We must make the optimum (not maximum) use of statistics, and must judge efficiency not by the volume, but by the extent to which they serve a distinct and clear purpose'.[2]

Market Research

The term 'market research' covers a wide range of activities which have a common object. That is to learn as much as possible about the people who buy the various products, why they buy them and what can be done to persuade them to buy more. Even the most enthusiastic market research consultant will concede that some firms are too eager to rush into this field and

[1] For this purpose the H.M.S.O. publication 'Economic Trends' and the bi-monthly bulletin of the National Institute of Economic and Social Research are useful.

[2] C. Davenport Hughes, 'Difficulties and Dangers in Statistics in Industry'. *The Incorporated Statistician*, March 1954.

spend money before finding out whether or not a market survey is necessary to solve their particular problem. Any reputable agency will invariably extract via 'desk research' the maximum information about the company's affairs before undertaking the survey. This has two advantages. It reveals whether or not the survey is really needed; for the cause of the malaise in the business may lie elsewhere. Furthermore, 'desk research' often reveals just what the problem is, so that the expenditure on market research may then be concentrated on that aspect of the firm's marketing policy which appears to fall short of requirements.

Many market research consultants complain that some firms regard their function in much the same light as many people regard their doctor. As long as all seems to be going on fairly well, they don't worry. It is only when business starts declining that they call in the services of the market research agency and then, on the strength of one survey – often done as cheaply as possible – expect them to effect a cure. There is a large body of expert opinion which considers that the advantages of market research are only fully derived from continuous market surveys. They do not deny that a single *ad hoc* survey designed to elucidate a particular problem may be very useful. They emphasise, however, the value for any firm to be kept abreast of current developments. For example, that firm's sales may appear to be well maintained but if market demand is growing it follows that the firm is then only retaining a declining portion of the market.

Some research agencies maintain consumer panels which continuously test and comment on various products. Since the cost of such a panel is probably spread over a number of clients, its 'services' can usually be bought for rather less than a full scale *ad hoc* sample survey. Similarly, the services of an organisation such as A. C. Nielsen of Oxford, which specialises in the well-known retail audit enquiry, are invaluable to any large producer of branded products. This provides a two-monthly statement of sales, current stocks held by retailers and changes therein, of a wide range of branded consumer goods, so that any one manufacturer can judge how his own product is faring in competition with other brands.

Even with a well-managed company which maintains its own statistical unit, recourse to the occasional market survey can be

highly profitable. The specialist agency has access to much information relative to many lines of business; their expertise and knowledge can yield results the value of which will repay many times the cost of their services. No reputable agency will undertake a survey for the mere sake of the fee. It knows full well that only if it can produce useful results will its client be pleased and wish to re-engage their services. So extensive has the use of market research become in industry today, not merely in the field of consumer purchases, but also in the field of capital goods such as plant and machinery, that there is much truth in the comment, 'the question is not whether we can afford it; it is whether we can afford to do without it!'

The Z Chart

A type of graph which enjoys a considerable vogue in business rather than in statistical circles, is that known as the Z chart. It derives its name from the form made by the lines on the graph, as will be seen from a scrutiny of Figure 25. The Z chart is merely a method of graphing a time series in such a way that the totals for successive periods are plotted and in addition the cumulative total and a moving annual total.

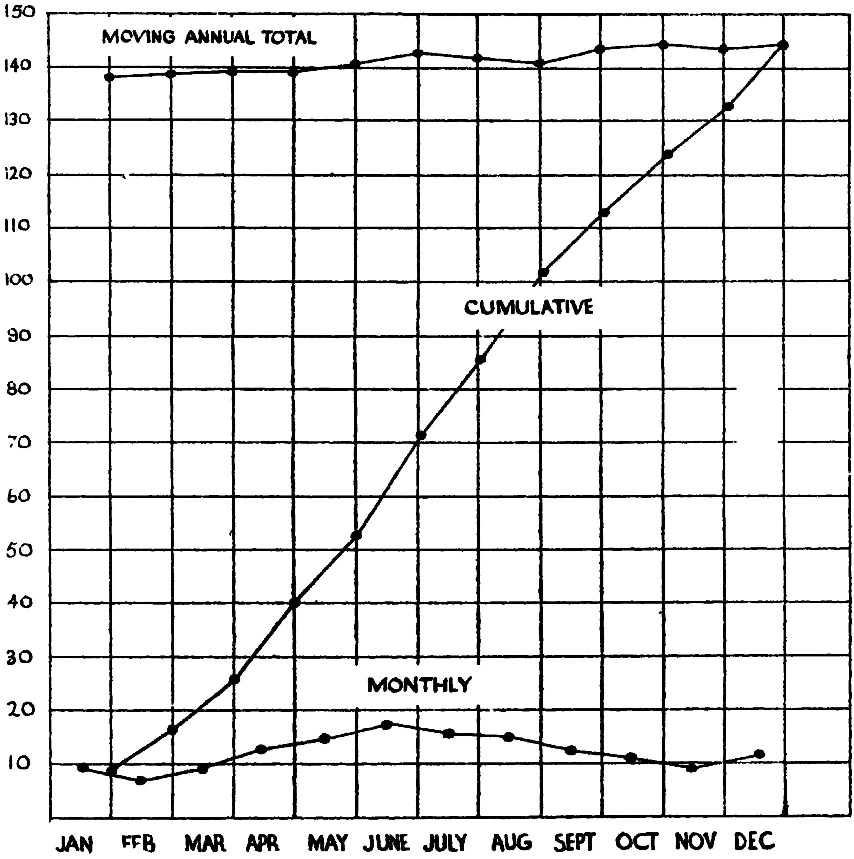
ABC Company Ltd. Sales Record, 1958

Month	Monthly Sales	Cumulative Monthly Total	Moving Annual Total
January.. ..	9,378	9,378	138,680
February ..	7,624	17,002	138,827
March	9,310	26,312	138,965
April	12,851	39,163	139,633
May ..	14,394	53,557	140,172
June ..	17,839	71,396	142,619
July	15,674	87,070	142,206
August	15,301	102,371	141,977
September ..	12,219	114,869	143,869
October ..	10,046	124,636	144,705
November ..	8,917	133,553	144,147
December ..	11,463	145,016	145,016

The data on which the graph is based are given above in the table, showing the sales of ABC Co., Ltd. This is divided into three columns. The first gives the monthly turnover, and the second the cumulative total as from January. The final column provides the annual total of sales for the twelve months

Figure 25

Z CHART



Sales record of ABC Company Limited 1958

ended in any month of the current year. Thus against June the figure in the third column is £142, 619, *i.e.*, the total sales for the twelve months ended 30th June 1950. The total for the period ended 31st July is £142,206, which is smaller than the preceding figure by £413. Since the sales for the current July were £15,301, the sales for the July in the previous year were £413 greater, *i.e.*, £15,714. The figures for all the months for the preceding year can be so computed if necessary from the table, except for January.

The real value of the moving annual total is that it indicates the trend of sales relatively to the preceding year's experience. If the moving annual total line is rising, it indicates that each month this year is an improvement on the same month of the preceding year. If required, a series of such charts for successive years can be set side by side for comparative purposes.

Business Forecasting

No branch of the statistician's broad field of enquiry provokes such controversy as that known as 'economic forecasting'. The fact is, however, that in business affairs the firm which can make the best estimate of future trends will make the largest profit. The government is nowadays much exercised to maintain a high and stable level of employment with increasing production. To these ends, both industry and the government have availed themselves of the services of statisticians. A recent publication by the Market Research Society entitled *Business Forecasting*[1] reveals that the leading firms in this country incur considerable expense in maintaining a statistical section which devotes a large part of its time to forecasting the future course of the business in which they are directly concerned. Similarly, the government has given its blessing to the work of the National Institute of Economic and Social Research in the field of economic forecasting.[2]

There are many types of forecast ranging from the 'hunch' of the business man that prices are shortly going to rise, to the efforts of the statisticians at the N.I.E.S.R. and Central Statistical Office to forecast the short run development of the economy by econometric methods based on a wide range of data of varying degrees of reliability and up-to-dateness. Only a fool would make extravagant claims for such 'guesstimates' as they have been called, but there have been a number of occasions when predictions were demonstrably accurate.[3] As pointed out above, the stakes are too high for any potentially useful tool to be ignored. If 'forecasting' can be of any assistance in ensuring that firms are not caught napping, then it will be undertaken. It remains therefore to consider the various techniques.

[1] Publications of the Market Research Society No.3.

[2] See the bi-monthly bulletin of the N.I.E.S.R.

[3] See 'Econometric and Sample Survey Methods of Forecasting' by L. R. Klein in *Market Research Society's* vol. *op. cit.*

The first method is to assemble data relating to the period just past, *e.g.* of sales, and to assume that any discernible trend will continue into the immediate future. In the absence of any other information this is quite a useful method, but it rests on the validity of the assumption that the conditions which determined the trend and pattern of sales during the period just elapsed, will continue to operate in the immediate future. An alternative method using the same data is to examine the factors and underlying forces which have produced those results. Once they have been ascertained it is up to the forecaster to decide whether those same factors are going to remain equally effective in the future or, if they are likely to change, how they will affect his firm. There is little question that in this field the statistician can do a good deal. The tremendous expansion in recent years in the volume of official statistics makes it possible for him to correlate one series with others and determine from the various series of data available the extent to which one factor rather than another has exerted any influence. For such work the statistician needs to have more than a passing familiarity with official statistics. While it may not be immediately obvious, the fact remains that many series are still of dubious value, *e.g.* those on stocks; while others appear only after a lengthy interval of time and may be out of date and therefore irrelevant to the problem.

The largest gap in the official statistics is that of what are called 'forward looking statistics'. These are derived from surveys of consumers' and industrialists' expectations and intentions regarding the future. In the United States such surveys are frequent and the government has expressed its satisfaction with the results.[1] In the United Kingdom these enquiries are still in their infancy. One of the best known is the Board of Trade's sample survey among large industrial undertakings to assess their intentions in the field of capital expenditure.[2] It is probably too early as yet to evaluate the reliability of these data but it is known that in one or two cases the government statisticians have rejected the results of the enquiry, preferring to rely upon their own estimates. It is noteworthy that the authors of another paper in the Market Research Society's book on market research in the firm of J. Lucas (Electrical) Ltd. commented that 'we

[1] See 'The Use of Official Statistics for Business Forecasting' by C. T. Saunders, *Market Research Society's* vol. *op. cit.*

[2] See p. 358.

found by past experience that we are not able to base future projections of our requirements upon consolidated manufacturers' plans. Manufacturers tend to be incurably optimistic and tend to talk in terms of maximum capacity only'.[1]

This may, of course, be a reaction peculiar to the circumstances of the post-war situation in the motor-car industry. On the other hand, those statisticians who advocate expanding this particular type of statistic would argue that no one would or should rely upon the actual figures themselves. Interest in these series lies primarily in their guide to the timing and the direction of change in the relevant variables.

The most recent and interesting development in economic forecasting is in the field of the econometric model. This is a system of simultaneous equations which is so designed as to represent the workings of the national economy. The 'model' is especially valuable when the statistician is dealing with variables which are prone to sudden change, *i.e.* in market situations where uncertainty is great. Its merit is that it explains why and how the changes in the economy or market are to take place; the forecast merely shows where the firm is going without specifying in detail the determining factors. The larger models may have nearly fifty equations based upon various data reflecting the relationship between variables such as the level of investment and consumer demand, imports and internal economic activity. This highly specialised field is the undisputed preserve of the expert and only begins to be comprehensible to the mathematician who requires the services of modern electronic computers to resolve these complex models. Nevertheless, it is not so much the solution of the equation which provides the real problem; it is that each equation has to be separately prepared in the light of the available data and each estimate of a relationship is a potential source of error.

The Statistician in Industry

The value of the statistician as a member of the production manager's staff will often be reflected in tangible evidence of lower production wastage with reduced inspection costs, etc. In this field the professional statistician's specialised mathematical skills are of especial importance. Probably the most widely

[1] *Application of Market Investigation*, by C. P. D. Davidson and J. T. Joyce.

known of all statistical techniques in industry of all kinds is *quality control.* As the name suggests, the object of this technique is to control the quality of a given product in such a way that wastage is minimised because it is detected as soon as scrap begins to appear.[1]

Every production manager has experienced the teething troubles with a prototype. Such experimentation is expensive and wasteful. Sometimes the product has to be modified over and over again until it meets the designer's requirements; on other occasions – and these are far fewer, it comes out right very quickly. A statistical approach via *experimental design* has been evolved whereby a specially designed series of experiments provides information which enables the statistician to determine which particular courses of action or development of a project are likely to yield useful results and those which are likely to be of little value. In other words, work can be concentrated upon those lines which are most likely to yield results, thereby saving both time and money. It is freely conceded by 'practical' production men that such analytical techniques are not merely labour savers, but may at times produce a solution to particularly complex problems which might never have been reached by conventional experimental techniques. A simple illustration is provided by the effort to produce a certain compound capable of resisting heat and abrasion up to certain limits. If there are several variables, which individually react upon each other, then the variations possible in the final product are very high indeed. Properly designed experiments can greatly reduce the number of attempts to achieve the final product of the required standard.

Operations Research

The wartime technique of *operations research* has since been adopted in some large organisations whereby teams of experts, with varying scientific but usually mathematical training, seek a solution to a particular problem in theoretical terms, and they then translate their theories into practice. Operations research is not purely statistical; but without the utilisation of modern statistical theory and practice, O.R. would not be the promising tool that it undoubtedly has become. Statistics to the O.R. team is what the scalpel is to the surgeon. Knowledge of the laws of

[1] See Chapter XX.

probability upon which statistical theory is based makes it possible to derive optimum solutions to what are known as 'queueing problems'. For example, how many berths should be built to discharge cargo from ships which arrive at varying intervals in such a way that berths are neither left unused for long periods, nor are so few in number that ships are kept waiting for long periods to discharge their cargoes. The same technique may be used in the solution of bottlenecks in the production line, or any service establishment, *e.g.* the canteen queue. This technique is at present one of the most flexible and powerful methods of analysis available. Similar techniques make it possible to determine the rate at which material stock should be ordered in a factory to ensure that the minimum amount of capital and space are tied up in holding such stock, and yet ensure that the chances of a production hold-up on account of stocks running out are kept to an absolute minimum.

In some complicated situations in which chance factors will often disturb a pre-arranged plan or schedule, the *Monte Carlo technique* is employed, but this is practicable only with electronic computers. In simple terms, this technique reproduces a great range of operational data for given sets of circumstances in a very short space of time. In this way the most effective counteraction can be prepared to meet the more common causes of disruption. Without it, management would only acquire such knowledge over a very long period with greater consequent wastage from frequent breakdowns. Another technique known as *linear programming* is beginning to attract much attention, although it is still in its early stages of development as a practical tool. Visualise, for example, a factory meeting a markedly seasonal demand. In one quarter demand greatly outstrips production, in another the reverse occurs. One solution is to produce for stock in the slack period; another is to employ intensive overtime in periods of peak demand. Each of these 'solutions' is expensive, the one because it ties up capital, the other because it is uneconomic to operate. Given the data relating to the limits of storage capacity, costs of overtime production, costs of annual output and a firm estimate of the prospective demand, 'linear programming' techniques permit a solution to this problem whereby for a given output total costs including warehousing, capital tied up, overtime charges, etc., are minimised.

Then there is the new *theory of games*. This is the study of competition between opponents in which a certain strategy is worked out by one player, the probable reactions of the opponent anticipated and the best counter devised. During the last war the technique was extensively applied in the air and sea war. The analysis is basically simple since it consists of nothing more than ascertaining the probable consequences of any course of action by one of the players so that the relative merits of each counter can be judged. As long as the strategy is restricted to two players with limited possibilities of action, the analysis is not difficult. It becomes extremely complex when the technique is applied to real life situations in view of the large number of variables. Hence the employment of computers is necessary.

Conclusions

It will be apparent from the foregoing illustrations that statistical techniques can find extensive application in the field of commerce and industry. While the simpler forms of analysis requiring 'desk research' demand relatively little mathematical knowledge, for business forecasting a familiarity with some mathematical techniques such as a curve fitting and multiple correlation is desirable. An understanding of the new and growing technique known as 'operations research' is restricted to the mathematician. Nevertheless, even a limited understanding of statistical techniques can help the administrator to appreciate what the professional statistician and the O.R. team are endeavouring to do. If full benefit from such techniques is to be obtained within the business world, there must be understanding between the technicians and the executive staff.

REFERENCES

There is, quite apart from the references given in the text, an expanding literature on the application of statistical methods in commerce and industry. The quarterly *Journal of the Association of Incorporated Statisticians* regularly contains articles discussing the structure of statistical units in different industries and their functions. The journal *Applied Statistics* of the Royal Statistical Society also contains readable and non-mathematical articles on the application of statistical methods to specific problems.

As was stated in the text, the various aspects of Operations Research are solely for the mathematical statistician. But the general reader who would like to learn more about the scope of the subject will find much to interest him, much of it in

non-mathematical language, in *Operations Research for Management* by McCloskey and Trefethen, published by the John Hopkins University Press, U.S.A. On *Econometrics*, Dr. Jan Tinbergen's book of that title, published by Allen and Unwin, provides an interesting introduction for the economist-statistician.

CHAPTER XX

QUALITY CONTROL

Introduction

One of the most useful yet simple applications of statistical theory based on the Normal curve is to be found in industry. The feature of modern industry is repetition work or mass production. The manufactured products may be intended for use by themselves, *e.g.*, rolls of cloth; or for use in conjunction with other parts made elsewhere, *e.g.*, a component of a machine. For all the precision of modern engineering, no two pieces following one another off the same machine are identical. The differences may be so small as to be invisible to the naked eye, but they exist. The simplest example will be appreciated by the reader of crime fiction. There is apparently no difficulty in determining from which of any number of apparently identical firearms a particular bullet was fired. The markings on the bullet are caused by the imperfections in the rifling, since the perfect boring instrument has yet to be devised. But this doesn't affect the efficiency of the firearm. Not all engineering products require such precision in manufacture and it is usual on an engineering blue-print to read machining tolerances. These tolerances indicate the permissible variation from the manufacturing specification. Such tolerances are necessary because variations in the product are inevitable; their sole purpose is to lay down limits which the variations should not exceed. One method widely employed to ensure that defective or inferior quality products are not passed into stock from the workshop is to have an inspection department. Usually the inspection is 100 per cent, every product being examined and the worker being paid on the accepted output. This system has two main weaknesses. The faulty work is detected only after it has been done and if several processes have been carried out after the piece became faulty, the machine and labour time wasted is considerable. Futhermore, human nature is such that even the 100 per cent inspection system is no guarantee that only satisfactory products will leave the works. Lastly, the cost of such inspection departments is often considerable.

The Theory of Quality Control

The best inspection system is that which detects the fault as soon as it appears, *i.e.*, at its origin, while also dispensing with 100 per cent manual inspection of the end product by substituting a virtually foolproof system of continuous sample inspection. Such a system is provided by the technique known as *Quality Control*. Since it is conceded that variation in the quality or size of a product is inevitable and within certain limits permissible, the first step is to ascertain the causes of any variation. Some variation in quality is certainly attributable to *chance*. This is usually so small and insignificant that it may be ignored; in any case since it is caused by numerous independent factors, it would be uneconomic and even impossible to trace them all. For example, the quality of brass castings being turned on a lathe will certainly vary, or the running speed of the machine may fluctuate. The important and larger variations are attributable to *assignable* causes. These are defects in the production process which by themselves will adversely affect the quality of the product, *e.g.*, excessive wear on the cutting tool, bad handling of the machine by the operative, and so on. These causes can and must be traced immediately their presence becomes apparent if the product is to be of the required standard.

Quality control enables those in charge of production to verify whether variations in the quality of the product are attributable to chance or to assignable causes. If they are of the latter type, then remedial action by the manager is called for. The function of the engineer at the commencement of production is to set the machine so that it may be expected to produce the particular product to the required specification. When this is done, 'controlled' conditions of production are created. Quality control indicates immediately to the engineer when production conditions cease to be 'controlled'. To illustrate general practice, a simple example is drawn from the light engineering industry, where quality control has been widely adopted. Assume that a hole is to be drilled in a particular product to the depth of 1·00 inches. The maximum tolerance permissible in the depth is, say, three-thousandths of an inch, *i.e.*, no hole may exceed 1·003 inches, or be less than ·997 inch. We may imagine that the drilling machine has been set and a special jig prepared to help the machinist to work fast, yet maintain the required standard. The

TABLE 91

Original Data. (All Figures in Inches)

Sample	1	2	3	4	5	6	7	8	9	10	11
	·999	1·003	·995	·996	·995	·996	·996	1·002	1·003	1·004	·996
	1·000	·997	·998	·998	1·000	1·001	1·005	·996	1·003	·999	1·000
	·999	·999	1·000	·996	·995	·995	·998	1·000	1·000	·996	1·003
	·998	1·000	·999	·995	·997	1·005	1·000	·998	1·002	1·003	·998
	1·004	1·005	1·000	·997	·997	·999	·997	1·000	·996	1·002	1·001
	·999	1·001	·999	1·005	1·002	1 000	·996	1·003	1·003	1·000	·998
	·999	·996	·998	1·004	·997	·999	1·002	1·005	1·000	1·004	·998
	1·004	1·005	·997	·999	1·003	·997	1·004	1·002	1·003	1·004	1·000
Total	8·002	8·006	7·986	7 990	7·986	7·992	7·998	8 006	8·010	8·012	7·994
Mean	1·00025	1·00075	·99825	·99875	·99825	·99900	·99975	1·00075	1·00125	1·00150	0·99925
Range	·006	·009	·005	·010	·008	·010	·009	·009	·007	·008	·007

Sample	12	13	14	15	16	17	18	19	20	Mean	Total
	1·002	1·004	·998	1·002	·997	1·004	1·003	1·001	·998		
	1·004	1·005	·997	·995	·996	1·003	1·001	·998	·996		
	·997	·998	1·002	1·005	1·000	·995	·996	·996	·996		
	1·000	1·004	1·002	·996	1·000	·996	·996	1·001	1·000		
	1·003	·999	1·002	·999	1·005	·995	1·002	·995	·993		
	1·002	1·004	·997	·999	1·004	1·005	·995	1·000	·995		
	·999	·999	1·002	1·002	·995	·998	·995	·998	·990		
	1 003	·997	1·004	1·002	·995	1·002	·996	·997	·992		
Total	8·010	8·010	8 004	8·000	7·992	7·998	7·984	7·986	7·960		159·926
Mean	1·00125	1·00125	1·00050	1·00000	·99900	·99975	·99800	·99825	·99500	·99955	
Range	·007	·008	·007	·010	·010	·010	·008	·006	·010	·0082	·164

Process Average = ·99954″ Mean Sample Range = ·0082″

next stage is to ascertain whether production conditions are controlled. For this purpose the product has to be examined. Either every unit drilled may be tested for depth, or samples may be taken at regular intervals. Let it be assumed that eight units are tested every half-hour, representing 10 per cent of the hourly output of 160. The proportion tested usually lies between 5 and 10 per cent of the total output, while individual samples may contain any number of units. The usual practice is for the inspector to take the last successive eight units off the machine on his arrival. Sometimes the sample may be selected at random by the inspector from the bin containing the finished product since his previous visit. Both methods have their advantages, but from the point of view of controlling the process the former method has more to recommend it. The dimensions of each of the eight units are set down as in Table 91, and the sample mean and range calculated. This is done for, say, 20 samples, and then the sample means and ranges are themselves aggregated and averaged. The resultants are termed the *process average* and the *mean sample range*, usually represented by the symbols $\bar{X}$ and w. They are, of course, estimates of the 'population' mean and dispersion for all the units produced under controlled conditions.

This assumption is based on the belief that as long as the process is controlled, all the drillings constitute a single homogeneous population. This follows since the only variations arise from chance and not from assignable causes; just like the distribution of heads in the tossing of sets of coins. If this is true, the population of drillings may be assumed to be normally distributed. It follows then that if chance alone causes variations, *i.e.*, the process remains controlled, only 1 in 20 units will vary beyond the limits of two standard deviations about the mean. If the limits are raised to 3·09 standard deviations, then it may be expected that only 2 in 1,000 will diverge by more than this. It need hardly be added that instead of calculating the range for the samples, the standard deviation might equally well have been derived. But it is so much easier to extract the range for small samples and it can be converted into terms of standard deviations by using special conversion tables. These give in most cases a good (*i.e.* satisfactory) conversion; but not a perfect conversion.

Control by Charts

Quality control is based on a continuous checking of the dimension under control, so that any variation greater and more frequent than that which may be expected from chance causes alone is immediately detected. Corrective action to the machine process is thereupon undertaken before serious waste ensues. Indication that action from the engineer is called for is provided by the *control charts*. These charts are based on the expectation that successive sample means drawn from a single homogeneous population will distribute themselves normally. Fig. 19 illustrates the simple form of combined chart which is extensively used. It comprises two charts, one above the other, so that they have a common X axis. The upper chart is for the sample averages, the lower one for sample ranges. For the present we are concerned only with the chart for sample averages.

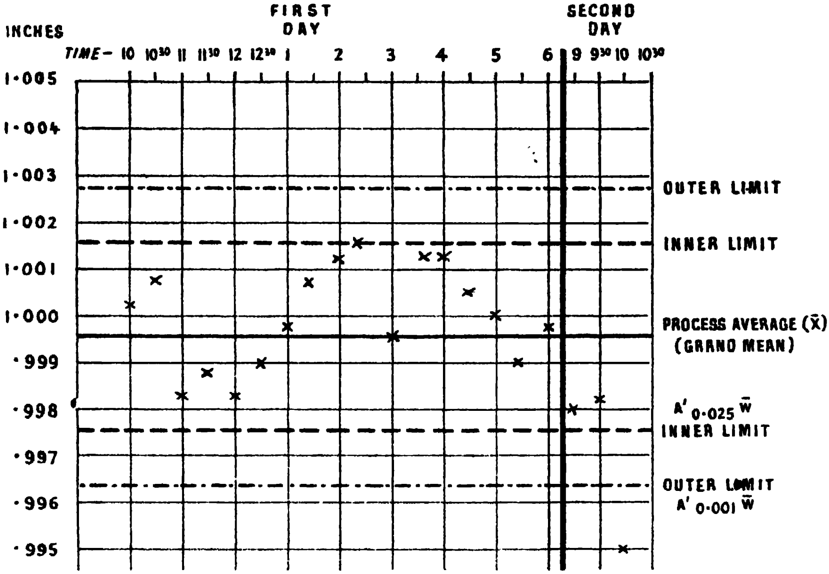
A line parallel to the base is drawn from a point along the Y axis, at ·99954 inch, equal to the process average derived from the data in Table 91. The base is marked off in half-hourly intervals. As the successive sample means are derived they are plotted on this chart. If the process is under control they should distribute themselves about the horizontal line representing the process average. To what extent may deviations from this line be tolerated? The permissible limits are indicated by the parallel lines drawn either side of the process average. The nearest pair are known as the Inner *control limits* and the outside pair as the Outer limits. The distances from the process average are equal to 1·96 and 3·09 standard deviations respectively; they are, in effect, the 5 per cent and 0·2 per cent confidence limits.

The values of the process average and the range are based initially on the data provided by a limited number of samples, *e.g.*, the 20 illustrated in Table 91. As the operation is continued and further data are collected, the original estimates can be revised where necessary.

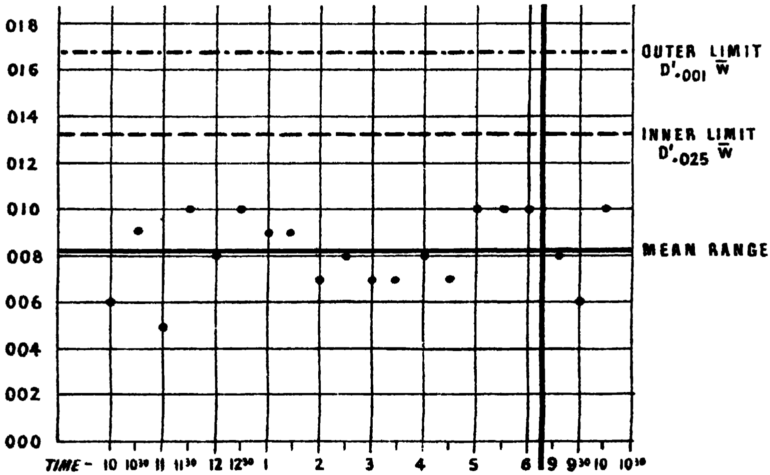
Reference has already been made to the conversion factors used to reduce the sample ranges to standard deviations. Special tables have been compiled for this purpose. These factors vary with the number of units in the sample and the extent of the sample range. This is to be expected, since we know that the value of the standard deviation is inversely related to the sample size. The conversion factors for deriving the control limits about

Figure 26

MEAN CHART



RANGE CHART



the process average from the sample ranges are represented by the letter A′.[1] To distinguish between the factors used for the Outer and Inner limits, the former is written $A'_{0\cdot001}$ and the other $A'_{0\cdot025}$. These are then multiplied by the mean sample range $\bar{w}$, so that the inner control lines are given by $\bar{\bar{X}} \pm A'_{0\cdot025}\ \bar{w}$ and the outer limits by $\bar{\bar{X}} \pm A'_{0\cdot001}\ \bar{w}$. Substituting for these symbols with the values given in Table 91, we get the inner control limits at ·9975 inch and 1·0015 inches from ·99954 ± ·244 (·0082); the outer control limits at ·9964 inch and 1·0027 inches from ·99954 ± ·384 (·0082).

The control chart for *ranges* is derived in similar fashion. In this case, too, the conversion factor is multiplied by $\bar{w}$ to derive inner and outer control limits representing the same probabilities as before. The factor in this case is symbolised by the letter D. The inner limit then is given by the expression $D'_{0\cdot025}$ w; substituting we get 2·04 (·0082) = ·01673 inch. The other limit is at ·01328, $D'_{0\cdot001}$ being equal to 1·62. Only two instead of four limits are given here and reference to the scale will reveal that the range is limited downwards by zero. The main purpose of this control chart is to see that ranges do not get out of control, *i.e.*, they do not exceed the mean range by more than a given amount. But if they approach zero this only implies that the production process is becoming even more stable and accurate than expected. This may, of course, be significant, but it can be ignored here. The importance of the range chart is probably clear. We know that the mean of any two samples can be the same, although the ranges are very different. As long as the deviations from the mean given by the two extreme values balance with regard to signs, they may become very large without affecting the mean. Returning to the illustration, if only the sample means were plotted, the engineer would not know if all units in the sample fell within the control limits, even though the sample mean did. As long as the production process is controlled the dispersion within samples will be kept fairly constant But, as will be seen, the first evidence of lack of control is often a greater dispersion of the sample values, although the means appear to be under control. It will be understood that the two charts are inter-related and should be read together.

[1] The values for A and D used in the following paragraphs are taken from Tables 10 and 13a of 'Quality Control Charts', by B. P. Dudding and W. J. Jennett, published by the British Standards Institution. This is the clearest authoritative description of British practice, and should be read by anyone interested in the subject.

Production Control by Charts

Once the control charts have been prepared the procedure is quite simple. The inspector, equipped with suitable gauges, checks the required sample as it comes off the machine at the specified intervals. Each dimension is recorded, the sample mean and range computed and plotted on the charts. As long as the plots oscillate regularly about their mean line, even if they occasionally drift towards either of the outer limits, the process is under control. As soon as any plot falls outside, remedial action is called for.

From the account so far, it would appear that for 10 hours the time apparently needed to collect the first 20 samples for preparing the charts, there is no continuous control. Only the units sampled are checked. In practice, however, the control charts are usually prepared after, say, 10 samples have been drawn. To avoid undue delay, the samples are often taken at shorter irregular intervals. From the data so derived preliminary charts can be drawn. After another twenty or so samples have been obtained at the usual interval, the data so obtained are used to revise, if necessary, the earlier charts. Clearly, a great deal will depend on the type of process to be controlled. Some can be set up quite easily; with others, charts cannot be prepared with any confidence for quite a long time.

To simplify the exposition, the data in Table 91 will be used for two purposes. They have already served to compute the control limits (for this purpose the last sample would have been excluded in practice, but the error can be ignored here).[1] Now let it be assumed that these data are being collected sometime after and represent the course of current production. For the first eight samples the product appears to be under control. The next two sample means show a tendency to stay around the upper limit and perhaps a small adjustment is made. The next series of samples in consequence drift across the graph, while the last sample falls well outside the outer limit. At this point a check would be called for and the units produced immediately before the sample was drawn should be checked.

Interpreting the Charts

Reading the charts is a straightforward matter, the interpre-

[1] Since the last sample is clearly faulty, it does not belong to the 'population'. In calculating the 'population' mean, this sample would normally be excluded.

tation is slightly more complex. There are two main types of excessive variation which will reveal their presence as the process goes out of control. The first, already mentioned, is that the sample means will tend to either of the limits and remain in close proximity to it. Even if the plots remain on the right side of the limit, the process should be checked. A simple example was illustrated above. The other fault is for the ranges to get out of control. In the above example, the charts show that the ranges are well under control, but very often when they do go out of control the trouble may be more difficult than with the means. The actual interpretation of the charts in terms of mechanical faults is the responsibility of the engineer, although experience of the process will often indicate its cause. When sample means tend to keep away from the process average, bad tool-setting or careless adjustment may be the cause. Increasing sample ranges may be attributable to excessive tool wear or undue play in worn bearings.

All that has been done so far is to establish the inherent nature of the process. On that information certain positive action can be based in connection with the production of the article. In practice, there is much more even to the simpler systems of quality control than this description may suggest. The sole purpose of discussing this technique is to emphasise the variety of applications of statistical theory. Briefly, however, the control charts should *ensure* that the final product will comply with the manufacturing specifications. One problem is to ensure that the chart control limits do not exceed the machining tolerances allowed in the design. In other words, when the specified machining limits lie well inside the range of $\pm 3\sigma$, then the process will produce more rejects than if the limits were wider. The solution in such a case is to examine whether these fine machine tolerances are necessary and if not to increase them. If they are, then the manufacturing process must be improved.

Advantages of Quality Control

The technique described above provides a virtually continuous inspection of the product. It can be applied at all stages of the manufacturing process, *e.g.*, there may be several drilling and milling operations. Each can be controlled independently. The outstanding advantages are as follows:

1. Quality control dispenses with the need for 100 per cent inspection of the finished product now so widely employed in industry. Human nature being what it is, this system is not 100 per cent reliable and quality control is probably more efficient and cheaper.

2. An outstanding feature is the reduction of wasted time and material to an absolute minimum since faulty production is immediately detected and the causes removed. An operator can see the likelihood of scrap being produced on his machine by observing the pattern of plots on his chart and can take action to avoid it. The saving in terms of labour, machine hours, material and supervision, quite apart from the reduction or even abolition of the final inspection department, more than cover the costs of introducing and maintaining a comprehensive quality control system.

3. The experience of industry with this technique has been such as to encourage many progressive firms to introduce it. Although specially suited to the light engineering industry, where it is often called 'dimensional' quality control, the principles are the same for other industries. Systems have been evolved for the chemical and textile industries among others. An interesting application of quality control technique was described in *Target*.[1] The sugar firm of Tate & Lyle Ltd packs weekly 7,500 tons of sugar into one and two pound packets. No machine can weigh this number of packets exactly and a margin – in favour of the consumer – must be allowed. By using quality control the average overweight can be regulated and in 18 months, using relatively unskilled female labour, the average overweight per packet has been reduced from $\frac{1}{2}$ to $\frac{1}{4}$ of a dram. This apparently insignificant saving represents 65 tons of sugar per annum. Further, unlike so many schemes introduced by management these days to increase production, workers soon appreciate the benefits of the system. The basis and operation of quality control can be explained quite simply to semi-skilled operatives and machinists on the floor of the shop. The female labour operating the system in the Tate & Lyle packing plants are merely given a brief training course. It also

[1] July 1953, Journal published by the Central Office of Information.

permits a better distribution of skilled supervisory capacity.

4. Combined with quality control, special sampling systems have been devised which provide a manufacturer with almost complete security against the risk of despatching inferior products to his customers. These methods are based on statistical theory beyond the scope of this text.

Conclusions

In its simplest form quality control represents an application of the normal curve theory and sampling. While sampling in the ordinary meaning of the term permits an approximate assessment of the whole, with statistical methods it enables one to judge whether or not successive small samples are of a different quality from each other. Given the knowledge acquired from a succession of samples, it is then possible to infer with considerable accuracy not merely the quality of the whole, but to forecast accurately the standard of quality which can be maintained in the future. Only the principles have been indicated here. There is, however, a growing literature on this subject, much of which is designed for the technical worker and non-statistician. Further reading will indicate some of the many applications of this useful technique and it is certain that further developments will follow rapidly as quality control is more widely employed.

APPENDIX A

MECHANISED PUNCHED CARD SYSTEMS

In several chapters reference was made to the use of punched cards in connection with survey work. It will be remembered that the questionnaire reproduced on pages 238 to 240 contained what were described as pre-coded answers to several of the questions. It was there suggested that the major virtue of this method was that it saved time during the interview and led to greater accuracy, since the interviewer did not have to waste time setting down verbatim, or in note form, the answer given by the respondent. On this score alone, pre-coding of answers on the schedule of questions is decidely useful. The main reason for the practice, however, is to be found in the widespread use of mechanical aids to sorting and tabulating the data collected on the interviewers' schedules. The various mechanical aids employed to this end are simply described as mechanised 'Punched Card' systems, due to the fact that the two I.C.T. systems,[1] the Hollerith and the Powers-Samas, are both based on the use of a punched card, which is described below. In passing, it may be said that broadly speaking, both systems achieve the same results and with few exceptions the main functions of the individual machines used in the two systems are identical. The principal difference lies in their method of functioning, for whereas with the Hollerith the sensing of the cards is by electrical contact through the holes punched in them, with Powers-Samas the sensing is mechanical by means of pins passing through the holes. 'Sensing' in this context means selecting and sorting the card according to the position of the punched hole in the cards.

The basis of the systems, it may be repeated, is the thin pasteboard card, approximately 7½in by 3in in size. In passing, it may be mentioned that cards are available with from 21 to 160 columns. The layout of the cards is specially designed for each particular survey or any other work which the machines will undertake. In each of the columns are series of figures. Some cards contain all the values from 0 to 9 inclusive and possibly at the head of the column two other code letters, *e.g.*, X and Y, if extra codes are needed. Thus if there are 12 answers possible to a particular question, all the values 0 to 9 and say X and Y are inserted in the column. When the schedule is returned by the interviewer, the marked number denoting the answer given, is punched on the card in that column. If necessary, combinations of any values may be used. The ruling between the various columns of figures breaks the cards up into what are known as 'fields'. These fields are headed by the title of that particular section of the schedule, and the relevant answers are punched within them. These mechanised punched card systems have other uses apart from statistical survey work. In actual fact, both systems find their main employment in the accounting field, although as social and commercial survey work expands, relatively more use will be made of the machines in this field.

[1] International Computers and Tabulators Ltd.

It may sometimes be necessary to use more than one card for each schedule. This has certain disadvantages which, however, can be quite easily overcome. For the moment, the point to bear in mind is the necessity for giving every care and attention to the layout of the card. This is a highly skilled technique, and I.C.T. Ltd. willingly help users of their equipment to design new layouts. The reason for this emphasis on layout is that without careful planning from the start, many of the virtues of this method, *e.g.*, speed, may be partially lost. For example, a poor layout may necessitate several sortings of many thousands of cards, when with a well designed card which had anticipated this particular analysis, two or three sortings would have sufficed.

However, assuming the cards to have been properly designed, the next stage is to transfer the information on the schedules to these cards. The punching of the cards is effected either by a small manually operated machine which resembles part of a typewriter keyboard; or by an automatic machine described at the end of this Appendix. The operator has merely to punch the code number ringed by the interviewer on to the same number in the appropriate field of the card. A competent operator can reach and maintain a remarkably high speed of punching. The main factor in determining the rate of punching, however, is usually the state of the schedule and the legibility of the writing! Since any statistics produced from these punched cards can be no more accurate than those punched on the card, it is essential to verify the accuracy of the operator's work.

Two methods of verification are provided. In the first, the cards are 're-punched' from the schedules by another operator. The machine used does not 'punch' but senses conformity with the holes already made. If either the first operator has made a mistake or the second operator accidentally depresses the wrong key, the machine locks. The operator must then check the cause of the disagreement. With the alternative method check punching is also used, but in this case mechanical 'sensing' is used to detect any disagreement between the first and second punching operations.

Once all the fields are punched, the card then constitutes a permanent record of all the information originally contained in the interviewer's schedule. Since the fields are headed, the operators can usually interpret quite easily the holes on the card, if a single card is being examined for any reason. It is at this stage that the value of this system becomes apparent, *i.e.*, sorting and tabulation. It will be remembered that so far the extraction of the information from the schedules is no more than is done manually on prepared forms.

The Sorting machine is capable of sensing the classification codes, *i.e.*, the holes, punched into the cards and arranging the cards into any classes or sub-classes, as the operator may pre-determine. Apart from sorting at speeds up to 40,000 cards an hour, the sorter will also record the number of cards in each class. Once sorted, the information must be printed in tabular form. This function is performed by the Tabulator, which not only prints on specially prepared forms any of the required information on the cards, but will also accumulate and print sub-totals and totals at pre-determined intervals. The machine may be set so as to add or subtract one set of group

totals to or from another, and so on. These machines are so flexible that they can be adjusted to reproduce any part or any combination of data contained in the cards, at speeds ranging from 5,000 to 12,000 cards per hour. The information contained in these sheets is then available for detailed study.

It can scarcely be denied that the first sight of these machines in operation is startling, in particular the speed of the sorter and the tabulator. But, invaluable as are all the functions described so far, the major advantage of mechanised systems has hardly been sufficiently stressed. It has already been pointed out that manual tabulation may be almost as rapid as the transfer of data to the punched cards; and the clerks then have the tables virtually prepared. But in all surveys, the respondents are classified into various groups and by various characteristics. Take, for example the Hulton Readership Survey on the results of which was based the study 'The Pattern of British Life'.[1] Manual tabulation will provide from the schedule such information as 'how many women smoke, how many read a women's weekly journal, how many have two children, live in a flat or a house, like electrical appliances and, if so, which do they own?' But if after sorting out all these results, the investigator wants to know for example, 'how many women living in flats with children have electrical appliances and what are the most popular ones', the only way to derive this information is by working all through the schedules again noting each woman to whom all these characteristics apply. With punched cards, assuming all this information has been punched, the task is easy. If the card has been designed with this type of analysis in mind, the sorting machine can select from all the cards, those with these particular characteristics. As many as eight characteristics can be selected simultaneously. This example has been deliberately exaggerated, but the basic principle remains true, *i.e.*, that for cross analyses of the various characteristics of the individuals sampled, the mechanised punched card systems are supreme. It was pointed out early in the appendix, that a great deal of unnecessary sorting and time can be saved by designing the cards in anticipation of the various analyses of the sample that may be required. But even if successive sortings of the cards, selecting one characteristic at a time, were necessary, it would still be faster than the manual method.

OTHER MACHINES

Automatic punching equipment has been referred to in an earlier paragraph of this appendix, and as it is possible that one or other of the various types available may be found useful, a very brief description of the more usual of their special functions is appended, together with brief details of other punched card machines which may also, in particular circumstances, prove of interest.

Automatic Punches

Several types of automatic punches are available, each embodying some or all of the following features:

[1] 'The Pattern of British Life'. Hulton Press Ltd.

(*a*) Reproducing punched data from one card on to another, or from one set of cards on to a corresponding set; or

(*b*) Gang punching data from a Leader, or Master Card on to any number of following cards; or

(*c*) Summary Punching or Balance Punching. In this operation the Automatic Punch is linked to the Tabulator, and set up to punch a card recording the information accumulated in the counters of the Tabulator, together with such indicative information as may be required, as each group, or sub-group of primary records is dealt with by the Tabulator; or

(*d*) Comparing the data punched into one set of cards with that punched into another corresponding set and automatically indicating any difference that may be sensed; or

(*e*) Mark Sense punching of manually marked cards. With this latest development of the automatic punch primary records may be entered direct on to cards, the entry being effected manually with a graphite pencil. These graphite markings will be subsequently sensed by the automatic punch and translated into punched holes. This sense punching is carried out at a rate of 6,000 cards per hour. It will be evident that where it is practicable to make primary entries on blank cards this method has much to recommend it.

Electronic Calculators

With the advent of electronics many difficulties which attended multiplying in the past – converting the factors into shillings and decimals or pence and decimals – have now disappeared. Modern punched card calculators can carry out a sequence of such operations as multiplying, cross adding, extracting of square roots, subtracting and dividing without converting any of the factors into decimals.

They can at the same time check punched calculations for correctness.

The Interpreter

As the name implies this machine is used for the purpose of interpreting cards, *i.e.*, sensing the data punched into a card and printing this information in alphabetical or numerical characters on the same card, usually along the top edge.

For certain sorting work which involves the intermingling of two sets of punched cards, a machine known as the Collator or Interpolator is available which will, at one simultaneous passage of two sets of cards through the machine intermingle them in correct numerical sequence, thus effecting considerable economy of sorting time. Whilst amalgamating the two packs, the machine will, if required, segregate from either pack all cards for which there are no corresponding cards in the other pack.

In addition to the compilation of statistics, the equipment described in the foregoing pages may also be employed to perform such day to day accounting procedures as Ledger Posting; Preparation of Payroll – including P.A.Y.E., Computation; Preparation of month-end Statements, etc., etc.

APPENDIX B

SOME SUGGESTIONS FOR FURTHER READING

This text, as was indicated in the Preface, is designed primarily for students studying part-time for their professional qualifications in which statistics is a subsidiary subject. Its scope is also sufficiently wide for those working for the external degrees of London University in the social sciences. Few part-time students find it possible to read extensively on any single subject, least of all a subsidiary one like statistics. Consequently a single book must usually serve. There is, however, always the possibility that sufficient interest in the subject may be aroused for the student to wish to pursue his studies further. With his limited background, the selection of suitable text from the very large number available is not easy, and the following suggestions may serve as a guide.

For the student who has no mathematics beyond the arithmetic in this text, the best book is *Principles of Medical Statistics*, by Prof. Bradford Hill. The adjective 'medical' should not be allowed to put off the prospective reader. It provides a brief but simple exposition of the principles of sampling and the various significance tests, some beyond the scope of the present book. In addition, it introduces the reader to some elementary ideas in the problems of statistical inference. For the reader who wishes to read more about the logic of significance tests and the basis of modern statistical analysis, 'An Introduction to Statistical Science in Agriculture' by D. J. Finney can be recommended. As with the first book mentioned, the title should not deter the reader. For the student interested especially in the application of statistical methods to economic data, a reading of selected chapters of *Applied General Statistics* by Croxton and Cowden, particularly those on index number theory and time series analysis, is instructive. The reader must be prepared to work through the numerous but brief examples which illustrate points made in these texts, but none of them is difficult. The student who possesses at least 'Advanced' level mathematics and who wishes to obtain a real understanding of statistical theory, must read *An Introduction to the Theory of Statistics*, by Yule and Kendall. While many sections of this authoritative work require but little mathematics, it can in no sense of the term be described as an 'introduction' for the reader ill-equipped with mathematics.

Occasional references (all of which are comprehensible to the non-mathematician) have been given in the text to reading on specific topics, but some of the following books may be of interest. On quality control there is a large body of literature, both British and American. Much of it is intended for technicians, but G. Herdan's *Quality Control* provides a detailed account which amalgamates the theory and practical aspects of this subject in simple language. *Statistical Methods in Research and Production*, edited by O. Davies, is a collection of concisely written essays on the application of the more advanced techniques in industry. Few proofs are given

of the theorems, and it is therefore much easier to read than the work by Yule and Kendall. On the other hand, the essays are all closely reasoned, and a fair knowledge of algebra is essential. But the reader with a special interest in this field who is prepared to take the subject seriously, this book will amply repay the time spent on it.

On sampling theory and procedures, the leading British book is *Sampling for Censuses and Surveys*, by Dr F. Yates. Some of this is descriptive in character, mainly those sections dealing with the conduct of surveys. But the major part, devoted to a discussion of sampling theory, is advanced. For the general reader *Methods of Social Surveys*, by C. A. Moser is the most suitable British book on Survey techniques. It contains extensive illustrations from current surveys. Apart from the occasional sections in the chapters on sampling, the book is eminently suited to the non-mathematical reader with an interest in this very important branch of statistical technique. A valuable feature is the very full bibliography of writings on survey techniques.

Whichever of the above-mentioned works is consulted, numerous references will be found to literature in that branch of statistics in which the reader wishes to specialise, always assuming, however, that he is prepared to devote more time to the study of the mathematical bases of statistics. Lastly, in the field of commercial statistics, the reader will note the growing use of graphs and statistical tables in the chairmen's reports which accompany the accounts of leading public companies. *The Annual Industrial and Financial Surveys* published by the *Financial Times*, *The Times* and the *Manchester Guardian* all contain articles which illustrate the use of national statistics in interpreting economic trends. Such sources provide invaluable lessons in the *art* of using statistics.

INDEX